

SCORE Search Results Details for Application 10552515 and Search Result us-10-552-515-1.rup.

[Score Home Page](#)
[Retrieve Application List](#)
[SCORE System Overview](#)
[SCORE FAQ](#)
[Comments / Suggestions](#)

This page gives you Search Results detail for the Application 10552515 and Search Result us-10-552-515-1.rup.

[start](#)

[Go Back to previous page](#)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2006, 20:20:01 ; Search time 318 Seconds
(without alignments)
2713.961 Million cell updates/sec

Title: US-10-552-515-1
Perfect score: 4950
Sequence: 1 MRMAATAWAGLQGPPLPTLC.....SELSSHWTPTVTPKASQLQQ 933

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_7.2:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	4950	100.0	933	2 Q6IWH7_HUMAN	Q6iwh7 homo sapien
2	3771.5	76.2	859	2 Q6IFT5_MOUSE	Q6ift5 mus musculu
3	3764	76.0	860	2 Q6IFT6_RAT	Q6ift6 rattus norv
4	1531.5	30.9	920	2 Q8NB39_HUMAN	Q8nb39 homo sapien
5	1525.5	30.8	955	2 Q32M45_HUMAN	Q32m45 homo sapien
6	1504	30.4	981	1 TM16C_HUMAN	Q9byt9 homo sapien
7	1498	30.3	956	2 Q6P5C6_MOUSE	Q6p5c6 mus musculu
8	1482	29.9	986	2 Q5XXA6_HUMAN	Q5xxa6 homo sapien
9	1479.5	29.9	840	2 Q8IYY8_HUMAN	Q8iyy8 homo sapien
10	1467.5	29.6	913	2 Q8CFW1_MOUSE	Q8cfw1 mus musculu
11	1464	29.6	999	1 TM16B_HUMAN	Q9nq90 homo sapien
12	1455	29.4	913	1 TM16E_HUMAN	Q75v66 homo sapien
13	1449	29.3	896	2 Q6DDQ3_XENLA	Q6ddq3 xenopus lae
14	1445	29.2	1219	2 Q9VTS0_DROME	Q9vts0 drosophila
15	1443	29.2	1235	2 Q2MOY5_DROPS	Q2m0y5 drosophila

16	1433	28.9	893	2	Q7QDY0_ANOGA	Q7qdy0 anopheles g
17	1402	28.3	904	1	TM16E_MOUSE	Q75ur0 mus musculu
18	1402	28.3	904	2	Q3V657_MOUSE	Q3v657 mus musculu
19	1396.5	28.2	910	1	TM16F_HUMAN	Q4kmq2 homo sapien
20	1384	28.0	1613	2	Q4S1F6_TETNG	Q4s1f6 tetraodon n
21	1383.5	27.9	871	2	Q8JFT1_BRARE	Q8jft1 brachydanio
22	1382.5	27.9	412	2	Q4SC97_TETNG	Q4sc97 tetraodon n
23	1371	27.7	911	1	TM16F_MOUSE	Q6p9j9 mus musculu
24	1369.5	27.7	1075	2	Q9VDV4_DROME	Q9vdrv4 drosophila
25	1367.5	27.6	712	2	Q8NCT7_HUMAN	Q8nct7 homo sapien
26	1367.5	27.6	926	2	Q8IN71_DROME	Q8in71 drosophila
27	1366.5	27.6	972	2	Q86P24_DROME	Q86p24 drosophila
28	1319	26.6	1035	2	Q4SSV5_TETNG	Q4ssv5 tetraodon n
29	1283.5	25.9	810	2	Q7QEP9_ANOGA	Q7qep9 anopheles g
30	1254.5	25.3	896	2	Q2VPA8_MOUSE	Q2vpa8 mus musculu
31	1249	25.2	984	2	Q8MT62_DROME	Q8mt62 drosophila
32	1237.5	25.0	971	2	Q4REV7_TETNG	Q4rev7 tetraodon n
33	1199.5	24.2	1043	2	Q9VYS8_DROME	Q9vys8 drosophila
34	1199.5	24.2	1059	2	Q76NS2_DROME	Q76ns2 drosophila
35	1154	23.3	596	2	Q8NB53_HUMAN	Q8nb53 homo sapien
36	1095.5	22.1	721	2	Q8AW48_BRARE	Q8aw48 brachydanio
37	1061.5	21.4	594	2	Q9NW72_HUMAN	Q9nw72 homo sapien
38	1058.5	21.4	594	2	Q8BHY3_MOUSE	Q8bhy3 mus musculu
39	1004.5	20.3	966	2	Q4RHQ9_TETNG	Q4rhq9 tetraodon n
40	1001	20.2	701	2	Q566T3_BRARE	Q566t3 brachydanio
41	912.5	18.4	475	2	Q8NAJ0_HUMAN	Q8naJ0 homo sapien
42	905	18.3	642	2	Q8N7V3_HUMAN	Q8n7v3 homo sapien
43	891.5	18.0	656	2	Q4SV37_TETNG	Q4sv37 tetraodon n
44	821	16.6	515	2	Q3VOR8_MOUSE	Q3vOr8 mus musculu
45	817.5	16.5	179	2	Q6IWH6_HUMAN	Q6iwh6 homo sapien

ALIGNMENTS

RESULT 1

Q6IWH7_HUMAN

ID Q6IWH7_HUMAN PRELIMINARY; PRT; 933 AA.
AC Q6IWH7;
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
DT 07-FEB-2006, entry version 10.
DE NGEF long variant.
GN Name=TMEM16G;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=14981236; DOI=10.1073/pnas.0308746101;
RA Bera T.K., Das S., Maeda H., Beers R., Wolfgang C.D., Kumar V.,
RA Hahn Y., Lee B., Pastan I.;
RT "NGEF, a gene encoding a membrane protein detected only in prostate
RT cancer and normal prostate.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:3059-3064 (2004).
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AY617079; AAT40139.1; -; mRNA.
DR HGNC; HGNC:31677; TMEM16G.
DR GO; GO:0005886; C:plasma membrane; IDA.
DR InterPro; IPR007632; DUF590.
DR InterPro; IPR002088; PPTA.
DR Pfam; PF04547; DUF590; 1.
DR PROSITE; PS00904; PPTA; UNKNOWN_1.
SQ SEQUENCE 933 AA; 105531 MW; D6FD42578A41D7D3 CRC64;

Query Match 100.0%; Score 4950; DB 2; Length 933;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 933; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRMAATAWAGLQGPPPLTLCPAVRTGLYCRDQAHAERWAMTSETSSGSHCARSRMLRRRA 60
|||||


```

OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=14981236; DOI=10.1073/pnas.0308746101;
RA Bera T.K., Das S., Maeda H., Beers R., Wolfgang C.D., Kumar V.,
RA Hahn Y., Lee B., Pastan I.;
RT "NGEP, a gene encoding a membrane protein detected only in prostate
RT cancer and normal prostate.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:3059-3064(2004).
CC -!- MISCELLANEOUS: The sequence shown here is derived from an
CC EMBL/GenBank/DBDJ third party annotation (TPA) entry.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; BK004075; DAA04566.1; -; mRNA.
DR Ensembl; ENSMUSG00000034107; Mus musculus.
DR MGI; MGI:3052714; Tmem16g.
DR InterPro; IPR007632; DUF590.
DR InterPro; IPR002088; PPTA.
DR Pfam; PF04547; DUF590; 1.
DR PROSITE; PS00904; PPTA; UNKNOWN_1.
SQ SEQUENCE 859 AA; 97128 MW; 82E1A473C59C8DA3 CRC64;

```

Qy	55	MLRRRAQEEDSTVLIDVS PPEAEKRGSYGSTAHASEPGGQQAACRAGSPAKPRIADFVL	114
Db	1	MLRGQAREEDSVVLIDMASPEAGNGCSYGSTAQASEAGKQQVAPSRVGSSAKPPI-DFVL	59
Qy	115	VWEEDLKLDRQQDASAARDRTDMHRTWRETFLDNLRAAGLCVDQDQDVQDGNTHVHYALLSA	174
Db	60	VWEEDL---RNQENPTKDKTDTHEVWRET FLENLCLAGLKLIDQHDVQDEAAAVHYILLRA	116
Qy	175	SWAVLCYYAEDRLRLKPLQELPNQASNWSAGLLAWLGINVLLLEVVPDPPEYYSCRFRV	234
Db	117	PWAVLCYYAEDRLRLKPLQELPNQASNWSATLLEWLGLPINILLEHVDPDPPEYYSCQFKA	176
Qy	235	NKLPRFLGSDNQDFTFTSTKRHQILFEILAKTPYGHEKKNLLGIHQLLAEGVLSAAFP LH	294
Db	177	SKLQWFLGSDNQDFTFTSTKRHQILFEILAKTPYGHEKGLFGIDQLLAEGVLSAAFP LH	236
Qy	295	DGPFKTPPEGPQAPRLNQOVLQFQHWARWGKWNKYQPLDHVRRYFGEKVALYFAWLGFYT	354
Db	237	DGPFSAVPESQVLGLIQOVLQFQHWARWGKWNKYQPLDHVRRYFGEKVALYFAWLGFYT	296
Qy	355	GWLLPAAVVGTLVFLVGCFLVFSDIPTQELCGSKDSFEMCPLCLDCPFWLLSSACALAAQ	414
Db	297	GWLLPAAVVGTVVFLVGCFLVFSDIPTQELCHSSDSFDMCPLCSDCSFWLLSSACTLAQA	356
Qy	415	GRLFDHGGTVFFSLFMALWAVLLEYWKRKSATLAYRWDCSDYEDTEERPRPQFAASAPM	474
Db	357	GRLFDHGGTVFFSLFMALWAVLLEYWKRKNATLAYRWDCSDYEDIERPRPQFAATAPM	416
Qy	475	TAPNPITGEDEPYFPPERSARRMLAGSVVIVVMVAVVMCLVSIILYRAIMAIIVSRSGN	534
Db	417	TALNPITGEDEPYFPEKNRVRRLAGSVVLLMMVAVVMCLVSVILYRAVMAIIVSRSDN	476
Qy	535	TLLAAWASRIASLTGSVVNLVLFILILSKIYVSLAHVLTREWEMHRTQTKFEDAFTLKVFIF	594
Db	477	AFLSAWASRIASLTGSVVNLVLFILILSKVYVLLAQVLTREWEMHRTQTEFEDAFTLKVFIF	536
Qy	595	QFVNFYSSPVYIAFFKGRFVGYPGNYHTLFGVRNEECAAGGCLIELAQELLVIMVGQVI	654
Db	537	QFVNFYASPVYIAFFKGRFVGYPGNYHTLFGIRNEECPAGGCLSELAQELLVIMVGQII	596
Qy	655	NNMQEVLIPKLKGWQKFLRLSKKRKAGASAGASQGPWEDDYELVPCGLFDEYLEMVLQ	714
Db	597	NNVQEVLPKLKGWQKF---SRGKKAG--TGTHPAPWEADYELLPCGLFHEYLEMVLQ	651
Qy	715	FGFVTIFVAAACPLAPLFAALLNNWEIRLDARKFVCEYRRPVAERAQDIGIWFHILAGLTH	774
Db	652	FGFVTIFVAAACPLAPLFAALLNNWEIRLDARKFVCEYRRPVAERAQDIGIWFHILGLTH	711

RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs.";
RL Nat. Genet. 36:40-45(2004).

Query Match 30.9%; Score 1531.5; DB 2; Length 920;
Best Local Similarity 37.9%; Pred. No. 5.7e-117;
Matches 360; Conservative 168; Mismatches 316; Indels 105; Gaps 29;

Qy 781 AFLALFSSDFLPRAYRW-----TRAHDLRGFLNFTLA-----RAP 816
 ||::| :||:| | : : ||:| :|:
 Db 751 AFVIAITSDFIPRLVYAYKYGPCAGQGEAGQKCMVGYNASLSVFRISDFENRSEPESDG 810
 Qy 817 SSFAAAHNRTCRYRAFRDDGH-----YSQTYWNLLAIRLAFVIVFEHVFSVGRLLDLL 871
 | | : ||| :| | :|:| | |||:||||:| : | : |
 Db 811 SEFSGTPLYKCRYRDYRDPHSLVPYGYTLQFWHVLAAARLAFIIVFEHLVFCIKHLISYL 870
 Qy 872 VPDIPESVEIKVKREYYLAKQALAENEVLFGTNGTKDEQPKGSELSSHW 920
 :||:| : :||:| | : : | | : | : | : |
 Db 871 IPDLPKDLRDRMRREKYLIQEMMYEALERLQKERKERKNGKAHHNEW 919

RESULT 5

Q32M45 HUMAN

ID Q32M45_HUMAN PRELIMINARY; . PRT; 955 AA.
 AC Q32M45;
 DT 06-DEC-2005, integrated into UniProtKB/TrEMBL.
 DT 06-DEC-2005, sequence version 1.
 DT 07-FEB-2006, entry version 3.
 DE TMEM16D protein.
 GN Name=TMEM16D;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
 OC Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=PCR rescued clones;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=PCR rescued clones;
 RG NIH MGC Project;
 RL Submitted (NOV-2005) to the EMBL/GenBank/DBJ databases.
 CC -----
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License
 CC -----
 DR EMBL; BC109308; AAI09309.1; -, mRNA.
 SQ SEQUENCE 955 AA; 111462 MW; 9A9348C61A4F20AF CRC64;

Query Match 30.8%; Score 1525.5; DB 2; Length 955;
 Best Local Similarity 38.2%; Pred. No. 1.9e-116;
 Matches 355; Conservative 163; Mismatches 309; Indels 103; Gaps 28;

Qy 63 EDSTVLID----VSPPEAE-----KRGSYGST---AHASEPGGQAAACRAGS-----PA 105
 :| :| | ||| : : | :| | ||| :
 Db 56 KDVNIFLDELEAVSSPCKDDSLHHPGNLTSTSDASRL EAGGETVPERNKSNGLYFRDG 115
 Qy 106 KPRIADFVLVWEEDLKLDRQQDS AARDRTDMHRTWRETFLDNLRAAGLCVDQDQDVGNT 165
 | || ||:| : : :| :| | ||| || : : : | :
 Db 116 KCRI-DYILVYRK-----SNPQTEK----REVFERNIRAEGLMQKESSLI-NS 158

Qy 166 TVHYALLSASWAVLCYYAEDLRLKLPLQE----LPNQASNWS-----AGLLAWLGIPNV 215
 Db 159 DII FVKLHAPWEVLGRYAEQMNVRMPFRRKIYYLPRRYKFMSRIDKQISRFRRLPKKPM 218

Qy 216 LL--EVVPDVP--EYYSRFRVKNLPRFLGSDNQDTFTSTKRHQILFEILAKTPYGHEK 272
 Db 219 RLDKETLPDLEENDCYTAPFSQQRHHFI-IHNKETFFNNATRSRIVHHILQRIKY-EEG 276

Qy 273 KNLLGIHQLLAEGVLSAAPFLHDGPFKTPPEGPQAPRLNQOVLFQHWARWGKWNKYQPL 332
 Db 277 KNKIGLNRLLTNGSYEAAFLHEGSYRSKNSIRTHGAENHRHLLYECWASWGVYKYQPL 336

Qy 333 DHVRRYFGEKVALYFAWLGFYTGWLLPAAVVGTLLVFLVGCFLVFSDIPTQELCGSKDSFE 392
 Db 337 DLVRRYFGEKIGLYFAWLGWYTGMLFPAAFIGLFVFLYGVTTLDHSQVSKEVCQATDII- 395

Qy 393 MCPLC-LDCPFWLLSSACALAQAGRLFDHGGTVFFSLFMALWAVLLLEYWKRKSATLAYR 451
 Db 396 MCPVCDKYCFMRLSDSCVYAKVTHLFDNGATVFFAVFMAVWATVFLEFWKRRRAVIAYD 455

Qy 452 WDCSDYEDTEERPRPQFAAS-APMTAPNPITGEDEPYFPERSRARRMLAGSVVIVVMVAV 510
 Db 456 WDLIDWEEEEIRPQFEAKYSKKERMNPISGKPEPYQAFTDKCSRLIVSASGIFFMICV 515

Qy 511 VVMCLVSIILYRAIMAVVSRSGNTLLA-AWA-----SRIASLTGSVV--NLVFILILSK 562
 Db 516 VIAAVFGIYIRVTV-----STFAAFKVALIRNNSQVAT-TGTAVCINFCCIIMLLNV 567

Qy 563 IYVSLAHVLRWEMHRTQTKFEDAFTLKVFIQFVNFYSSPVYIAFFKGRFVGYPGNYHT 622
 Db 568 LYEKVALLTNLQPRTESEWENSFTLKMFLFQFVNLNSSTFYIAFFLGRFTGHPGAYLR 627

Qy 623 LFG-VRNEECAAGGCLIELAQELLVIMVGKQVINNMQEVLPKLKGWWQKFLRSKKRKA 681
 Db 628 LINRWRLECHPSGCLIDLCMQMGIIMVLKQTWNNFELGYPLIQNWTR---RKVRQEH 684

Qy 682 GASAGASQGPWEDDYELVPCE--GLFDEYLEMVLQFGFVTIFVAACPLAPLALLNNWVE 739
 Db 685 GPERKISFPQWEKDYNLQPMNAYGLFDEYLEMILQFGFTTIFVAAFPLAPLALLNNIIE 744

Qy 740 IRLDARKEVCEYRRPVAERAQDIGIWFHILAGLTHLAVISNAFLAFSSDFLPRAYRW- 798
 Db 745 IRLDAYKFVTQWRRPLASRAKDIGIYGILEGIGILSVITNAFVIAITSDFIPRLVYAYK 804

Qy 799 -----TRAHDLRGFLNFTLA-----RAPSSFAAAHNRTCryRAFRDD 835
 Db 805 YGPCAGQGEAGQKCMVGYVNASLSVFRISDFENRSEPESDGSEFSGTPLKYCRYRDYRDP 864

Qy 836 DGH-----YSQTYWLLAIRLAFVIVFEHVVSFVGRLLDLLVPDIPESVEIKVKREYLA 890
 Db 865 PHSLVPYGYTLQFWHVLARLAFIIVFEHLVFCIKHLISYLIPDLPKDLRDRMRREKYLI 924

Qy 891 KQALAENEVLFGTNGTKDEQPKGSELSSHW 920
 Db 925 QEMMYEAELERLQKERKERKKNGKAHHNEW 954

RESULT 6

TM16C_HUMAN

ID TM16C_HUMAN STANDARD; PRT; 981 AA.

AC Q9BYT9;

DT 16-JAN-2004, integrated into UniProtKB/Swiss-Prot.

DT 01-JUN-2001, sequence version 1.

DT 07-FEB-2006, entry version 20.

DE Transmembrane protein 16C.

GN Name=TMEM16C; Synonyms=C11orf25;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;

OC Homo.

OX NCBI_TaxID=9606;

RN [1]

RP NUCLEOTIDE SEQUENCE [MRNA].

RA Rosier M.F., Toselli E., Segurens-Soury B., Auffray C., Devignes M.D.;

RT "Predominant brain expression and full-length characterization of a

novel human 6.6-Kb transcript mapping at 11p14 in the telomeric part


```

      ||||| ||| ::|||: || ||||| || |: ||||:||||:| :||:| | | :
Db      787 IRLDAYKFVTQWRRPLPARATDIGIWLGILEGIGILAVITNAFVIAITSFYIPRFVYEEK 846
Qy      799 -----TRAHDLRGFLNFTLARAP-SSFAAAHNRTCryRAFR-----DDGHYSQTY 843
      : ||:|:| :|: | : | | | :| : |
Db      847 YGPCANHVEPSENCLKGYVNNLSFFDLSELGMGKSGYCRYRDYRGPWSSKPYEFTLQY 906
Qy      844 WNLALIRLAFVIVFEHVVSFVGRLLDLLVPDIPESVEIKVKREYLLAKQALAENEV 899
      |:| | |||:|||||:| : : |:|:|: : :|| | | : : | | :
Db      907 WHILAAARLAFIIVFEHLVFGIKSFIAYLIPDVPKGLHDIRREKYLQVEMMYEAE 962

```

RESULT 7

Q6P5C6_MOUSE

```

ID      Q6P5C6_MOUSE  PRELIMINARY;  PRT;  956 AA.
AC      Q6P5C6;
DT      05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT      05-JUL-2004, sequence version 1.
DT      07-FEB-2006, entry version 12.
DE      AU040576 protein.
GN      Name=Tmem16a; Synonyms=AU040576;
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC      Muroidea; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RC      STRAIN=C57BL/6; TISSUE=Eye;
RX      MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA      Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA      Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA      Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA      Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA      Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA      Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA      Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA      Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA      Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA      Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA      Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA      Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA      Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA      Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA      Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA      Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA      Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT      "Generation and initial analysis of more than 15,000 full-length human
RT      and mouse cDNA sequences.";
RL      Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN      [2]
RP      NUCLEOTIDE SEQUENCE.
RC      STRAIN=C57BL/6; TISSUE=Eye;
RA      Strausberg R.;
RL      Submitted (NOV-2003) to the EMBL/GenBank/DDBJ databases.
CC      -----
CC      Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC      Distributed under the Creative Commons Attribution-NoDerivs License
CC      -----
DR      EMBL; BC062959; AAH62959.1; -; mRNA.
DR      Ensembl; ENSMUSG00000031075; Mus musculus.
DR      MGI; MGI:2142149; Tmem16a.
DR      GO; GO:0016021; C:integral to membrane; RCA.
DR      InterPro; IPR007632; DUF590.
DR      Pfam; PF04547; DUF590; 1.
SQ      SEQUENCE 956 AA; 110489 MW; 150FACCBDA4AF25 CRC64;

```

```

Query Match      30.3%; Score 1498; DB 2; Length 956;
Best Local Similarity 37.6%; Pred. No. 3.5e-114;
Matches 360; Conservative 171; Mismatches 303; Indels 124; Gaps 28;

```

```

Qy      26 GLYCRDQAHAEWAMT--SETSSGSHCARSRMLRRRAQEEDSTVLIDVSPPEAEKRGSYG 83
      ||| || : : : || | | | : | : |
Db      52 GLYFRDGKRKVDYILVYHHKRASG-----SRTLARRGLQNDMVL-----GTRS 94
Qy      84 STAHASEPGGQAAACRAGSPAKPRIADFLVWEEDLKLDRQQDSAARDRTDMHRTWRET 143

```

```

Db      95 VRQDQPLPG---KGSFVDAGSPEVP-----MDYHEDD-----KRFRREE 130
Qy      144 FLDNLRAAGLCVDQDQDVGNTTVH---YALLSASWAVLCYYAEDLRLKPLQELPNQAS 200
      : | | | | : : | : | : : | | | | | : | : | : : : :
Db      131 YEGNLEAGLELE---NDEDTKIHGVEFKIHAPWHVLCREAEFLKLMPTKKVYHISE 186
Qy      201 NWSAGLLAWLGIPNVLLEVVPDPPEYYSCRFRVNKLPRFLGS-----DNQDTFFT 251
      : | | | | : | : | : : | : | : | : | : | : | : | :
Db      187 --TRGLLK--TINSVLQKITDPIQPKVAEHRPQTTRLSYPFSREKQHLFDLTDRDSFFD 242
Qy      252 STKRHQILFEILAKTPYGHEKKNLGLIQLLAEGVLSAAFLHDGPFKTPPEGPQAPRLN 311
      | | | : | : | : | : | : | : | : | : | : | : | :
Db      243 SKTRSTIVYEILKRTTCTKAKYS-MGITSLANGVYSAAYPLHDGDY----EGDNV-EFN 296
Qy      312 QRQVLFQHWARWGKWNKYQPLDHRVRYFGEKVALYFAWLGFYTGWLLPAAVVGTLVFLVG 371
      | : | : | : | : | : | : | : | : | : | : | : | : | :
Db      297 DRKLLYEAWASYGVFYKQPIDLVRKYFGEKVGLYFAWLGYTQMLIPASIVGVIVFLYG 356
Qy      372 CFLVFSDIPTQELCGSKDSFEMCPLC-LDCPFWLLSSACALAQAGRLFDHGGTVFFSLFM 430
      | | | : | : | : | : | : | : | : | : | : | : | :
Db      357 CATVDENIPSMEMCDQRYNITMCLCDKTCYWKMSACATARASHLFDNPATVFFSVFM 416
Qy      431 ALWAVLLLEYWKRKSATLAYRWDCSDYEDTEERPRPQFAA-----SAPMTAPNPITGEDE 485
      | | | | : | : | : | : | : | : | : | : | : | : | :
Db      417 ALWAATFMEHWKRKQMRNLNYRWDLTGFEEEDHPRAEYEARVLEKSLRKESRNKET--DK 474
Qy      486 PYFPERSRARRMLAGSVVIVVMVAVVMCLVSIILYRAIMAVVSRSGNTLLAAWASRIA 545
      | | | : | : | : | : | : | : | : | : | : | : | :
Db      475 VKLTWRDRFPAYFTNLVSIIFMIAVTFIVLGVIIYRISTAAALAMNSSPSVRSNIRVTV 534
Qy      546 SLTGSVNVNLVFIILSKIYVSLAHVLTWRWEMHRTQTKFEDAFTLKVFIFQFVNFYSSPVY 605
      : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db      535 TATAVINLVVILLDEYVGCARWLTKIEVPKTEKSFEERLTFKAFLLKFVNSYTPIFY 594
Qy      606 IAFFKGRFVGYPGNYHTLF-GVRNEECAAGGCLIELAQELLVIMVGKQVI--NNMQEVLIP 663
      : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db      595 VAFFKGRFVGRPGDYVYIFRSFRMEECAPGGCLMELCIQLSIIMLGKQLIQNNLFEIGIP 654
Qy      664 KLKGWQKFRRLRSKKRKAGASAGASQGPWEDDYELVPCGLFDEYLEMVLQFGFVTIFVA 723
      | : | : : | : : : : | : | : | : | : | : | : | : | :
Db      655 KMKKFIRYLKLRQSPSDREEYVKKRQRYEVDNLEPFAGLTPEYMEMIIQFGFVTIFVA 714
Qy      724 ACPLAPLFAALLNNWVEIRLDARKFVCEYRRPVAERAQDIGIWFHILAGLTHLAVISNAFL 783
      : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db      715 SFPLAPLFAALLNNIIIEIRLDAKKFVTELRPVAIRAKDIGIWINLRGVGKLAVIINAFV 774
Qy      784 LAFSSDFLPRAYRWTRAHD--LRGFLNFTLARAPSSF-----AAAHN-----R 825
      : : | : | : | : : : : | : | : | : | : | : | : | :
Db      775 ISFTSDFIPRLVLYMYSQNGTMHGFVNHTL----SSFNVSDFQNGTAPNDPLDLGYEVQ 830
Qy      826 TCRYRAFRD---DDGHY--SQTYWNLALIRLAFVIVFEHVVSFVGRLLDLLVPDIPESVE 880
      : | : : | : | : | : | : | : | : | : | : | : | : | :
Db      831 ICYKDYREPPWSEHKYDISKDFWAVLAARLAFVIVFQNLVFMMSDFVDWVIPDIPKDIS 890
Qy      881 IKVKREYYL-----AKQALAENEVLFGTNGTKDEQPKGSELSSHWPFTVPKA 928
      : : | : | : | : | : : : | : : : | : | : | : | :
Db      891 QQIHKEKVLMLVELFMREEQGKQQLLDTWM-----EKEKPRDVPCNNH-SPTTHPEA 940

```

RESULT 8

Q5XXA6_HUMAN

ID Q5XXA6_HUMAN PRELIMINARY; PRT; 986 AA.

AC Q5XXA6;

DT 23-NOV-2004, integrated into UniProtKB/TrEMBL.

DT 23-NOV-2004, sequence version 1.

DT 07-FEB-2006, entry version 6.

DE Tumor amplified and overexpressed sequence 2.

GN Name=TMEM16A;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;

OC Homo.

OX NCBI_TaxID=9606;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RA Huang X., Godfrey T.E., Gollin S.M.;
 RT "Comprehensive Analysis of the 11q13 Amplicon in Oral Squamous Cell
 RT Carcinoma Cells and Synteny to Mouse Chromosome 7F5, a Band Amplified
 RT in Murine Oral Carcinoma.";
 RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
 CC -----
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License
 CC -----
 DR EMBL; AY728143; AAU82085.1; -; mRNA.
 DR HGNC; HGNC:21625; TMEM16A.
 DR InterPro; IPR007632; DUF590.
 DR Pfam; PF04547; DUF590; 1.
 SQ SEQUENCE 986 AA; 114078 MW; E30A02F91EF36FC2 CRC64;

Query Match 29.9%; Score 1482; DB 2; Length 986;
 Best Local Similarity 36.8%; Pred. No. 7.7e-113;
 Matches 365; Conservative 162; Mismatches 302; Indels 164; Gaps 30;

Qy	26	GLYCRDQAHAERWAMT--SETSSGSHCARSRLRRRAQEEDSTVLIDVSPPEAEKRGSYG	83
		: : : :	
Db	52	GLYFRDGRKVDYILVYHKKRPSG-----NRTLVRVQHS DTP-----SGA	92
Qy	84	STAHASEPGGQAAACRAGSPAKPRIADFVLVWEEDLKLDRQDQAARDRTDMHRTWRET	143
		: : : : : : : : : :	
Db	93	RSVKQDHPLPGKGASLDAGSGEPP-----MDYHEDD-----KRFREE	130
Qy	144	FLDNLRAAGLCVDQDQDGNNTTVH---YALLSASWAVLCYAE DLRLKLPLQELPNQAS	200
		: : : : : : : : : : : : : : :	
Db	131	YEGNLEAGLELE---RDEDTKIHGVGFKIHAPWNVLCREAEFLKLMPTKKMYH--I	184
Qy	201	NWSAGLLAWLGINVLLEVVPDPPEYYSCR-----FRVKNLPRFLGSDNQDTFF	250
		: : : : : : : : : : : : :	
Db	185	NETRGLLK--KINSVLQKITDPIQPKVAEHRPQTMKRLSYFSSREKQHLFDLSD-KDSFF	241
Qy	251	TSTKRHQILFEILAKTPYGHEKKNLLGIHQLLAEGVLSAAFPLHDGPFKTPPEGPQAPRL	310
		: : : : : : : : : : : : :	
Db	242	DSKTRSTIVYEILKRTTCTKAKYS-MGITSLLANGVYAAAYPLHDGDY-----NGENVEF	295
Qy	311	NQRQVLFQHWARWGKWNKYQPLDHVRRYFGEKVALYFAWLGFYTGWLLPAAVVGTLVFLV	370
		: : : : : : : : : : : : : : : :	
Db	296	NDRKLLYEWEARYGVFYKYQPIDLVRYFGEKIGLYFAWLGVYTMPLIPASIVGIIVFLY	355
Qy	371	GCFLVFSDIPTQELCGSKDSFEMCPLC-LDCPFWLLSSACALAQAGRLFDHGGTVVFFSLF	429
		: : : : : : : : : : : : : : :	
Db	356	GCATMDENIPSMEMCDQRHNITMCPLCDKTC SYWKMSACATARASHLFDNPATVFFSVF	415
Qy	430	MALWAVLLLEYWKRKSATLAYRWDCSDYEDTEE----RPRPQFAA-----SAPMTAPNPI	480
		: : : : : : : : : : : : :	
Db	416	MALWAATFMEHWKRKQMLNRYRWDLTGFEEEEEA VKDHPRAEYEARVLEKSLKKESRKN-	474
Qy	481	TGEDEPYFPERS-----RARRMLAG-----SVVIVVMVAVVVM	513
		: : : : : : : : : : : : :	
Db	475	--EKRRHIPEESTNKNWKQRVKTAMAGVKLT DKVKLTWRDRFPAYLTNLVSIIFMIAVTFA	532
Qy	514	CLVSIILYRAIMAIVVSRSGNTLLAASRIASLTG SVVNLVFIILSKIYVSLAHVLTR	573
		: : : : : : : : : : : : : : : : : :	
Db	533	IVLGVIIYRISMAAALAMNSSPSVRSNIRVTVTATAVIINLVVILLDEVYGC IARWLTK	592
Qy	574	WEMHRTQTKFEDAFTLKVFIFQFVNFIYSSPVYIAFFKGRFVGYPGNYHTLF-GVRNEECA	632
		: : : : : : : : : : : : : : : :	
Db	593	IEVPKTEKSFEERLIFKAFLLK FVNSYTPIFYVAFFKGRFVGRPGDYVYIFRSFRMEECA	652
Qy	633	AGGCLIELAQELLVIMVGKQVI--NNMQEVLIPKLGWQKFRLSKKRKAGASAGASQGP	691
		: : : : : : : : : : : : :	
Db	653	PGGCLMELCQLSIIMLGKQLIQNNLFEIGIPKMKKLIRYLKLKQQSPDHEECVKRKQR	712
Qy	692	WEDDYELVPCEGLFDEYLEMVLQFGFVTIFVAACPLAPL FALLNNWVEIRLDARKFVCEY	751
		: : : : : : : : : : : : :	
Db	713	YEVDYNLEPFAGLTPEYMEMIIQFGFVTLFVASFPLAPL FALLNNIEIRLDAKKFVTEL	772
Qy	752	RRPVAERAQDIGIWFHILAGLTHLAVISNAFLAFSSDFLPRA--YYRWTRAHDLRGFLN	809
		: : : : : : : : : : : : :	
Db	773	RRPVAVRAKDIGIWNILRGIGKLAVIINAFVISFSDFIPLRVLYLYMSKNGTMHGFVN	832
Qy	810	FTLARAPSSF-----AAAHN-----RTCRYRAFRD---DDGHY--SQTYWNLLA	848

```

Db      833 HTL-----SSFNVSDFQNGTAPNDPLDLGYEVQICRYKDYREPPWSENKYDISKDFWAVLA 888
Qy      849 IRLAFVIVFEHVVSFVGRLLDLLVPDIPESVEIKVKREYYLA-----KQALAENE 898
      |||||:::| : :|::|::: : :::| | || |
Db      889 ARLAFVIVFQNLVMEFMSDFVDWVDPDKDISQQIHKEKVLMMVELFMREEQDKQQLL--E 946
Qy      899 VLFGTNGTKDEQP-----KGSELSSH 919
      ||| | || |
Db      947 TWMEKERQKDEPPCNHHNTKACPDSLGSPPAPSH 979

```

RESULT 9

Q8IYY8_HUMAN

```

ID   Q8IYY8_HUMAN   PRELIMINARY;   PRT;   840 AA.
AC   Q8IYY8;
DT   01-MAR-2003, integrated into UniProtKB/TrEMBL.
DT   01-MAR-2004, sequence version 2.
DT   07-FEB-2006, entry version 13.
DE   TMEM16A protein.
GN   Name=TMEM16A;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Testis;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Testis;
RG   NIH MGC Project;
RL   Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
CC   -----
CC   Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC   Distributed under the Creative Commons Attribution-NoDerivs License
CC   -----
DR   EMBL; BC033036; AAH33036.2; -, mRNA.
DR   Ensembl; ENSG00000131620; Homo sapiens.
DR   InterPro; IPR007632; DUF590.
DR   Pfam; PF04547; DUF590; 1.
SQ   SEQUENCE 840 AA; 97654 MW; F8503B8F813CDA27 CRC64;

```

```

Query Match      29.9%;   Score 1479.5;   DB 2;   Length 840;
Best Local Similarity  40.0%;   Pred. No. 9.9e-113;
Matches 340; Conservative 152; Mismatches 270; Indels 89; Gaps 22;

```

```

Qy      135 DMHRTWRETFLDNLRAAGLCVDQDQVDGNTTVH---YALLSASWAVLCYYAEDLRLKLP 191
      | | || : || ||| :: :| :| :| : :| ||| || |::|
Db      6  DDRKFRREEYEGNLLEAGLELE----RDEDTKIHGVGFKIHAPWNVLCREAEFLKLKMP 61
Qy      192 LQELPNQASNWSAGLLAWLGIPNVLEVVPDVPPEYYSCR-----FRVNKLPRFL 241
      :: : | : || | :|| :: :| : | | |
Db      62 TKKMYH--INETRGLLK--KINSVLQKITDPIQPKVAEHRPQTMKRLSYFPSREKQHLFD 117

```

Qy 242 GSDNQDTFFSTKRHQILFEILAKTPYGHEKKNLLGIHQLLAEGVLSAAFPLHDGPFKTP 301
 Db 118 LSD-KDSFFDSKTRSTIVYEILKRTTCTKAKYS-MGITSLLANGVYAAAYPLHDGDY--- 172

Qy 302 PEGPQAPRLNQRQVLQFQHWARWGKWNKYQPLDHVRRYFGEKVALYFAWLGFYTGWLLPAA 361
 Db 173 --NGENVEFNDRLKLYEEWARYGVFYKYQPIDLVRKYFGEKIGLYFAWLGVYTQMLIPAS 230

Qy 362 VVGTLVFLVGCFLVFSDIPTQELCGSKDSFEMCPLC-LDCPFWLLSSACALAQAGRLFDH 420
 Db 231 IVGIIVFLVGCATMDENIPSMEMCDQRHNITMCPLCDKTCYWKMSACATARASHLFDN 290

Qy 421 GGTVFFSLFMALWAVLLLEYWKRKSATLAYRWDCSDYEDTEERPRPQFAA-----SAPMT 475
 Db 291 PATVFFSVFMALWAATFMEHWKRKQMRNLNRYWDLTGFEEDHPRAEYEAREVLEKSLKKE 350

Qy 476 APNPITGEDEPYFPERSRARRMLAGSVVIVVMVAVVMCLVSIILYRAIMAVVSRSGNT 535
 Db 351 SRNKET--DKVKLTWRDRFPAYLTNLVSIIFMIAVTFAIVLGVIIYRISMAAALAMNSSP 408

Qy 536 LLAAWASRIASLTGSVNVLVFILILSKIYVSLAHVLRWEMHRTQTKFEDAFTLKVFIFQ 595
 Db 409 SVRSNIRVTVTATAVIINLVVILLDEVYGCARWLTKIEVPKTEKSFEERLIFKAFLK 468

Qy 596 FVNFYSSPVYIAFFKGRFVGYPGNYHTLF-GVRNEECAAGGCLIELAQELLVIMVGKQVI 654
 Db 469 FVNSYTPIFYVAFFKGRFVGPRGDYVYIFRSFRMEECAPGGCLMELCIQLSIIMLGKQLI 528

Qy 655 -NNMQEVLIPKLKGWQKFRRLRSKKRKAGASAGASQGPWEDDYELVPCEGLFDEYLEMVL 713
 Db 529 QNNLFEIGIPKMKKLIRYLKLKQQSPPDHEECVKKRQRYEVDYNLEPFAGLTPEYMEMII 588

Qy 714 QFGFVTIFVAACPLAPLFAALLNNWVEIRLDARKFVCEYRRPVAERAQDIGIWFHILAGLT 773
 Db 589 QFGFVTLFVASFPLAPLFAALLNNIIIEIRLDARKFVTELRRPVAVRAKDIGIWNILRGIG 648

Qy 774 HLAVISNAFLFSSDFLPRA--YYRWTRAHDLRGFLNFTLARAPSSF-----AAAHN 824
 Db 649 KLAVIDAFVISFSDFIPLRVLYLYMYSKNGTMHGFVNHTL----SSFNVSDFQNGTAPN 704

Qy 825 -----RTCRYAFRD---DDGHY--SQTYWNLLAIRLAFVIVFEHVVSFVGRLLDL 870
 Db 705 DPLDLGYEVQICRYKDYREPPWSENKYDISKDFWAVLAARLAFVIVFQNLVFMFSDFVDW 764

Qy 871 LVPDIPESVEIKVKREYLA-----KQALAENEVLFGTNGTKDEQP----- 911
 Db 765 VIPDIPKDISQIHKEKVLMLVELFMREEQDKQQLL--ETCMEKERQKDEPPCNHHNTKAC 822

Qy 912 ---KGSELSSH 919
 Db 823 PDSLGSPAPSH 833

RESULT 10

Q8CFW1_MOUSE

ID Q8CFW1_MOUSE PRELIMINARY; PRT; 913 AA.
 AC Q8CFW1;
 DT 01-MAR-2003, integrated into UniProtKB/TrEMBL.
 DT 01-MAR-2003, sequence version 1.
 DT 07-FEB-2006, entry version 16.
 DE Transmembrane protein 16B.
 GN Name=Tmem16b;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muroidea; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Eye;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Eye;
 RA Strausberg R.;
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 CC -----
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License
 CC -----
 DR EMBL; BC033409; AAH33409.1; -; mRNA.
 DR Ensembl; ENSMUSG00000038115; Mus musculus.
 DR MGI; MGI:2387214; Tmem16b.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR InterPro; IPR007632; DUF590.
 DR Pfam; PF04547; DUF590; 1.
 KW Transmembrane.
 SQ SEQUENCE 913 AA; 104388 MW; CA17DB27D8167F64 CRC64;

Query Match 29.6%; Score 1467.5; DB 2; Length 913;
 Best Local Similarity 38.1%; Pred. No. 1.1e-111;
 Matches 331; Conservative 167; Mismatches 274; Indels 97; Gaps 22;

```

Qy      78 KRGS---YGSTAH--ASEPGGQQAACRAGSPAKPRIADFLVWEEDLKLDROQDSAARD 132
      ||: | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      20 KRGAHLGHGSPGHSLAVISNGETGKERHGGGPGDVELGLDALEEE----- 65

Qy      133 RTDMHRTWRETFLDNLRAAGLCVDQDQDQDNTTVHYALLSASWAVLCYYAEDLRKLPL 192
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      66 ----RREQRDEFHNLMAAGLEL-EKDLESKSQGSVFVRIHAPWQVLAREAEFLKIKVPT 120

Qy      193 QEL-----PNQASNWSAGLLAWLGIPNVLLVVPDPPEYYSCRFRVNKLP-----RF 240
      :: | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      121 KMYEIKAGGSIKKFSA-ILQTLSSP-----LQPRV-PEHSNNRMKNLSYPFSREKMYL 173

Qy      241 LGSDNQDTFTSTKRHQILFEILAKTPYGHEKKNLLGIHQLLAEGVLSAAFPLHDGPFKT 300
      : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      174 YNIQEKDTFFDNATRSRIVHEILKRTACS-RANNTMGINSLIANNIYEAAYPLHDGEYDS 232

Qy      301 PPEGPAQAPRLNQRQVLFQHWARWGKWNKYQPLDHVRRYFGEKVALYFAWLGFYTGWLLPA 360
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      233 PGDD-----MNDKLLYQEWARYGVFYKFPIDLRKYFGEKIGLYFAWLGLYTSFLIPS 287

Qy      361 AVVGTLVFLVGCFLVFSDIPTQELCGSKDSFEMCPLC-LDCPFWLLSSACALAQAGRLFD 419
      : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      288 SVIGVIVFLYGCATIEEDIPSKEMCDHQNAFTMCPLCDKSCDYWNLSSACGTARASHLFD 347

Qy      420 HGGTVFFSLFMALWAVLLEWYKRKSATLAYRWDCSDYEDTEER----PRPQF----- 468
      : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      348 NPATVFFSIFMALWATMFLENWKRLQMRGLGYFDLTGIEEEERSQEHRSPEYETKVREK 407

Qy      469 -----AASAPMTAPNPITGE-DEPYFPERSRARRMLAGSVVIVVMVAVVVMCLVSIIL 520
      : | : | : | | | | : | | | | | | | : | | : | : | : | : | : |
Db      408 LLKESGKSAVQKLEANSPEDEDEDKLTWKDRFPGYLMNFASILFMIALTFSIVFGVIV 467

Qy      521 YRAIMAIIVVSRSGNTLLAAWASRI---ASLTGSVVNLVFIILSKIYVSLAHVITRWEMH 577
      || | | : | | | | : | | | | | | | | | | | | | | | | | | |
Db      468 YRITTAALS-----LNKATRSNVRTVTATAVIINLVILDEIYGAVAKWLTKIEVP 522

Qy      578 RTQTKFEDAFTLKVFIFQFVNFYSSPVYIAFFKGRFVGYPGNYHTLF-GVRNEECAAGGC 636
      : | : | | | : | | | | | | | | | | | | | | | | | | | | | |
  
```

```

Db      523 KTEQTFEERLILKAFLLKFVNAYSPIFYVAFFKGRFVGRPGSYVYVFDGYRMEECAPGGC 582
Qy      637 LIELAQELLVIMVGKQVI--NNMQEVLIPKLKGWQKFRRLRSKKRKAGASAGASQGPWEDD 695
      | : | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      583 LMELCIQLSIIMLGKQLIQNNIFEIGVPKLKKLFRKLKDETEPGESDPDHSKRPEQWDLD 642
Qy      696 YELVPCEGLFDEYLEMVLQFGFVTIFVAACPLAPLFAALLNNWVEIRLDARKFVCEYRRPV 755
      : | | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      643 HSLEPYTGLTPEYMEMIIQFGFVTLFVASFLAPVFALLNNVIEVRLDAKKFVTELRRPD 702
Qy      756 AERAQDIGIWFHILAGLTHLAVISNAFLAFSSDFLPRAYRWTRAHD--LRGFLNFTLA 813
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db      703 AVRTKDIGIWFILSGIGKFSVIINAFVIAVTSDFIPRLVYQYSYSHNGTLHGFEVNHTLS 762
Qy      814 -----RAPSSFAAAHNRTCRYAFRD-----DDGHYSQTYWNLLAIRLAFVIVF 857
      : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db      763 FFNVSQLEKGTQFENSQFDQEVQFCRFKDYREPPWAPNPYEFKQYWSVLSARLAFVIIF 822
Qy      858 EHVVSFVGRLDLLVPDIPESVEIKVKRE 886
      : : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db      823 QNLVMFLSVLVDWMIPDIPTDISDQIKKE 851

```

RESULT 11

TM16B_HUMAN

```

ID      TM16B_HUMAN      STANDARD;      PRT;      999 AA.
AC      Q9NQ90;
DT      16-JAN-2004, integrated into UniProtKB/Swiss-Prot.
DT      01-OCT-2000, sequence version 1.
DT      07-FEB-2006, entry version 23.
DE      Transmembrane protein 16B.
GN      Name=TMEM16B; Synonyms=C12orf3;
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC      Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      NUCLEOTIDE SEQUENCE [MRNA].
RC      TISSUE=Retina;
RA      Lorenz B., White K.E., Econs M.J., Strom T.M.;
RT      "Transcripts in l2p13.3.";
RL      Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
CC      -!- SUBCELLULAR LOCATION: Membrane; multi-pass membrane protein
CC      (Probable).
CC      -!- SIMILARITY: Belongs to the TMEM16 family.
CC      -----
CC      Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC      Distributed under the Creative Commons Attribution-NoDerivs License
CC      -----
DR      EMBL; AJ272204; CAC01125.1; -; mRNA.
DR      Ensembl; ENSG00000047617; Homo sapiens.
DR      HGNC; HGNC:1183; TMEM16B.
DR      LinkHub; Q9NQ90; -.
DR      InterPro; IPR007632; DUF590.
DR      Pfam; PF04547; DUF590; 1.
KW      Membrane; Polymorphism; Transmembrane.
FT      CHAIN      1      999      Transmembrane protein 16B.
FT                                     /FTId=PRO_0000072564.
FT      TRANSMEM      360      382      Potential.
FT      TRANSMEM      535      557      Potential.
FT      TRANSMEM      577      599      Potential.
FT      TRANSMEM      619      641      Potential.
FT      TRANSMEM      746      768      Potential.
FT      TRANSMEM      796      818      Potential.
FT      TRANSMEM      898      920      Potential.
FT      VARIANT      108      108      V -> A (in dbSNP:3741903).
FT                                     /FTId=VAR_021932.
FT      VARIANT      501      501      S -> A (in dbSNP:1860961).
FT                                     /FTId=VAR_020331.
SQ      SEQUENCE      999 AA; 113616 MW; B9B4F56161AE1B00 CRC64;

```

Query Match 29.6%; Score 1464; DB 1; Length 999;
 Best Local Similarity 37.4%; Pred. No. 2.4e-111;
 Matches 344; Conservative 167; Mismatches 284; Indels 124; Gaps 27;

Qy 80 GSYGSTAHASEPGGQAAACRAGSPAKPRIADFLVWEEDLKLDRQQDSAARDRTDMHRT 139
 Db 125 GETGKEPHAGGPG-----DIELG-PLDALEERKEQ--- 154

Qy 140 WRETFLDNLRAAGLCVDQDQDGNNTTVHYALLSASWAVLCYYAEDRLRLPLQELPNQ- 198
 Db 155 -REEFEHNLMEAGLEL-EKDLENKSQGSIFVRIHAPQWVLAREAEFLKIKVPTKKEMYEI 212

Qy 199 -----ASNWSAGLLAWLGIPNVLLVVPDVPPEYSCRFRVNKLP-----RFLGSDNQ 246
 Db 213 KAGGSIKKFSAAL-----QKLSSHLQPRV-PEHSNNMKMNLSPFSREKMYLYNIQEK 265

Qy 247 DTFFTSTKRHQLFEILAKTPYGHEKKNLLGIHQLLAEGVLSAAFPLHDGPFKTPPEGPQ 306
 Db 266 DTFDNATRSRIVHEILKRTACS-RANNTMGINSLIANNIYEAAYPLHDGEYDSEDD-- 322

Qy 307 APRLNQRQVLFQHWARWGKWKYQPLDHVRRYFGEKVALYFAWLGFYTGWLLPAAVVGTL 366
 Db 323 ---MNRKLLYQEWARYGVFYKFQPIDLRKYFGEKIGLYFAWLGLYTSFLIPSSVIGVI 379

Qy 367 VFLVGCFLVFSDIPTQELCGSKDSFEMCPLC-LDCPFWLLSSACALAQAGRLFDHGGTVF 425
 Db 380 VFLYGCATIEEDIPSRMCDQQAFTMCPLCDKSCDYWNLSACGTAQASHLFDNPATVF 439

Qy 426 FSLFMALWAVLLLEYWKRKSATLAYRWDCSDYEDTEER-----PRPQFAA----- 470
 Db 440 FSIFMALWATMFLENWKRLQMLRGYFWDLTGIEEEERAEQHSRPEYETKVREKMLKESN 499

Qy 471 -SAPMTAPNPIT----GEDEPYFPERSRARRMLAGSVVIVVMVAVVMCLVSIILYRAIM 525
 Db 500 QSAVQKLENTTECGDEDEDKLTWKDRFPGYLMNFASILFMIALTFISIVFGVIVYRITT 559

Qy 526 AIVVSRSGNTLLAAWASRI---ASLTGSVNVNLFILILSKIYVSLAHVLTRWEMHRTQTK 582
 Db 560 AAALS-----LNKATRSNVRVTATAVIINLVVILILDEIYGAVAKWLTKIEVPKTEQT 614

Qy 583 FEDAFTLKVFIFQVNFYSSPVYIAFFKGRFVGYPGNYHTLF-GVRNEECAAGGLIELA 641
 Db 615 FEERLILKAFLLKFVNAYSPIFYVAFKGRFVGRPGSYVYVFDGYRMEECAPGGCLMELC 674

Qy 642 QELLVIMVGKQVI-NNMQEVLPKLGWQKFRRLRSKKRKAGASAGA-SQGP--WEDDYE 697
 Db 675 IQLSIIMLGKQLIQNNIFEIGVPKLK---KLFRKLKDETEAGETDSAHSKHPEQWDLDS 731

Qy 698 LVPCEGLFDEYLEMVLQFGFVTIFVAACPLAPLFAALLNNWVEIRLDARKFVCEYRRPVAE 757
 Db 732 LEPYTGLTPEYMEMIIQFGFVTLFVASFPLAPVFAALLNNVIEVRLDAKFKVTELRRPDAV 791

Qy 758 RAQDIGIWFHILAGLTHLAVISNAFLAFSSDFLPRAYRWTRAHD--LRGFLNFTLA-- 813
 Db 792 RTKDIGIWFIDILSGIGKFSVISNAFVIAITSDFIPRLVYQYSYSHNGTLHGFMNHTLSFF 851

Qy 814 -----RAPSSFAAHNRCTCRYAFRD-----DDGHYSQTYWNLLAIRLAFVIVFEH 859
 Db 852 NVSQLKEGTQPEQSQFDQEVQFCRFKDYREPPWAPNPYEFKQYWFILSARLAFVIFQN 911

Qy 860 VVFSVGRLLDLLVPDIPESVEIKVRE-----YYLAKQALAENEVLFGTNGTKDEQPKG 913
 Db 912 LVMFLSVLVDWMIPDIPTDISDQIKKEKSLLVDFFLKE----EHEKCLKMDEPALRSPGG 967

Qy 914 SELSSHWTPFTVPKA-SQL 931
 Db 968 GDRSRSRASSAPSGQSQL 986

RESULT 12

TM16E_HUMAN

ID TM16E_HUMAN STANDARD; PRT; 913 AA.

AC Q75V66;

DT 13-SEP-2005, integrated into UniProtKB/Swiss-Prot.

DT 05-JUL-2004, sequence version 1.

DT 07-FEB-2006, entry version 12.

DE Transmembrane protein 16E (Gnathodiaphyseal dysplasia 1 protein).

GN Name=TMEM16E; Synonyms=GDD1;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
 OC Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [MRNA], TISSUE SPECIFICITY, SUBCELLULAR LOCATION,
 RP VARIANTS GDD GLY-356 AND ARG-356, AND CHARACTERIZATION OF VARIANTS GDD
 RP GLY-356 AND ARG-356.
 RC TISSUE=Skeletal muscle;
 RX PubMed=15124103; DOI=10.1086/421527;
 RA Tsutsumi S., Kamata N., Vokes T.J., Maruoka Y., Nakakuki K.,
 RA Enomoto S., Omura K., Amagasa T., Nagayama M., Saito-Ohara F.,
 RA Inazawa J., Moritani M., Yamaoka T., Inoue H., Itakura M.;
 RT "The novel gene encoding a putative transmembrane protein is mutated
 RT in gnathodiaphyseal dysplasia (GDD).";
 RL Am. J. Hum. Genet. 74:1255-1261(2004).
 RN [2]
 RP TISSUE SPECIFICITY.
 RX PubMed=15067359;
 RA Katoh M., Katoh M.;
 RT "Identification and characterization of TMEM16E and TMEM16F genes in
 RT silico.";
 RL Int. J. Oncol. 24:1345-1349(2004).
 CC -!- SUBCELLULAR LOCATION: Membrane; multi-pass membrane protein
 CC (Probable). Endoplasmic reticulum. Co-localized with
 CC CALR/calreticulin.
 CC -!- TISSUE SPECIFICITY: Highly expressed in brain, heart, kidney,
 CC lung, and skeletal muscle. Weakly expressed in bone marrow, fetal
 CC liver, placenta, spleen, thymus, osteoblasts and periodontal
 CC ligament cells.
 CC -!- DISEASE: Defects in TMEM16E are the cause of gnathodiaphyseal
 CC dysplasia (GDD) (MIM:166260); also called osteogenesis imperfecta
 CC with unusual skeletal lesions or gnathodiaphyseal sclerosis. GDD
 CC is a rare skeletal syndrome characterized by bone fragility,
 CC sclerosis of tubular bones, and cemento-osseous lesions of the
 CC jawbone. Patients experience frequent bone fractures caused by
 CC trivial accidents in childhood; however the fractures healed
 CC normally without bone deformity. The jaw lesions replace the
 CC tooth-bearing segments of the maxilla and mandible with fibrous
 CC connective tissues, including various amounts of cementum-like
 CC calcified mass, sometimes causing facial deformities. Patients
 CC also have a propensity for jaw infection and often suffer from
 CC purulent osteomyelitis-like symptoms, such as swelling of and pus
 CC discharge from the gums, mobility of the teeth, insufficient
 CC healing after tooth extraction and exposure of the lesions into
 CC the oral cavity.
 CC -!- SIMILARITY: Belongs to the TMEM16 family.
 CC -----
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License
 CC -----
 DR EMBL; AB125267; BAD17859.1; -, mRNA.
 DR Ensembl; ENSG00000171714; Homo sapiens.
 DR HGNC; HGNC:27337; TMEM16E.
 DR MIM; 166260; phenotype.
 DR MIM; 608662; gene.
 DR InterPro; IPR007632; DUF590.
 DR Pfam; PF04547; DUF590; 1.
 KW Disease mutation; Endoplasmic reticulum; Glycoprotein; Membrane;
 KW Transmembrane.
 FT CHAIN 1 913 Transmembrane protein 16E.
 FT /FTId=PRO_0000191755.
 FT TOPO_DOM 1 299 Cytoplasmic (Potential).
 FT TRANSMEM 300 320 Potential.
 FT TOPO_DOM 321 380 Extracellular (Potential).
 FT TRANSMEM 381 401 Potential.
 FT TOPO_DOM 402 462 Cytoplasmic (Potential).
 FT TRANSMEM 463 483 Potential.
 FT TOPO_DOM 484 511 Extracellular (Potential).
 FT TRANSMEM 512 532 Potential.
 FT TOPO_DOM 533 557 Cytoplasmic (Potential).
 FT TRANSMEM 558 578 Potential.
 FT TOPO_DOM 579 679 Extracellular (Potential).
 FT TRANSMEM 680 700 Potential.
 FT TOPO_DOM 701 732 Cytoplasmic (Potential).
 FT TRANSMEM 733 753 Potential.
 FT TOPO_DOM 754 834 Extracellular (Potential).

Db 821 HNMQFWHVLAAKMTFIIIVMEHVFLVKFLLAWMIPDVPKDVVERIKREKLMTIKILHDFE 880
Qy 899 V 899
:
Db 881 L 881

RESULT 13

Q6DDQ3_XENLA

ID Q6DDQ3_XENLA PRELIMINARY; PRT; 896 AA.
AC Q6DDQ3;
DT 16-AUG-2004, integrated into UniProtKB/TrEMBL.
DT 16-AUG-2004, sequence version 1.
DT 07-FEB-2006, entry version 12.
DE Tmem16e-prov protein.
GN Name=tmem16e-prov;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipioidea; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Spleen;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Spleen;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Spleen;
RA Klein S., Gerhard D.S.;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; BC077486; AAH77486.1; -; mRNA.
DR InterPro; IPR007632; DUF590.
DR Pfam; PF04547; DUF590; 1.
SQ SEQUENCE 896 AA; 105167 MW; EBFCF828411C8C50 CRC64;

Query Match 29.3%; Score 1449; DB 2; Length 896;
Best Local Similarity 37.1%; Pred. No. 3.6e-110;
Matches 334; Conservative 167; Mismatches 297; Indels 102; Gaps 27;

Qy 59 RAQEEDSTVLIDVSPPEAEKRGSY---GSTAHASEPGGQQAACRAGSPAKPRIADFLV 115
| :|| ||::| | : | : || : | |
Db 7 RREEE---TLIEMSVTGDESNGALLDNNISITDSELPGNSEI-----DKHVQSKDSVFF 56

Qy 116 WEEDLKLD-----RQQDAAARDRTDMHRTWRETFLDNLRAAGLCVDQDQVQDG-N 164
 Db 57 WDGIRRIDFILSYTDETNKEAEKKAERRD-----FEFNLHKSGLELETEDKKDSEN 108

Qy 165 TTVHYALLSASWAVLCYAE DLRLKLPLQ--ELPNQASNWSAGLLAWLGIPNVLLLEVVPD 222
 Db 109 GKTYFLKIHAPWEVLTTTAEVLNIMKPLKADDLTDESENLLSHMLKPKFLP-----PE 161

Qy 223 V---PPEYYSRFRVNLPRFLGSDNQDTFFTSTKRHQILFEILAKTPYGHEK-KNLLGI 278
 Db 162 VMSPEPDYFTAPFRKDKQELFRIED-KEKFFTPSTRNRIVYIILSRCHYGEEGKKKFGI 220

Qy 279 HQLLAEGVLSAAFLHDGPF--KTPEGPQAPRLNQRVLFQHWARWGKWNKYQPLDHVR 336
 Db 221 KRLNNGSYLDAYPLHDCRYWKTD-----ERSCERYTLYSHWAKFTREYKEQPLDLIR 274

Qy 337 RYFGKEVALYFAWLGFYTGWLLPAAVVGTFLVFLVGCFLVFSDIPTQELC--GSKDSFEMC 394
 Db 275 KYYGEKIGIYFAWLGFYTEMLFYAAVVGFFCFLYGWITMDDSSISKEICDPGIGGQIIMC 334

Qy 395 PLC-LDCPFWLLSSACALAQAGRLFDHGGTVFFSLFMALWAVLLLEYWKRKSATLAYRWD 453
 Db 335 PLCDKRCEFWRLNSTCEPSQYSHMFDNVATLFFAIFMGIWVTLFLEFWKRRQARLEYEWD 394

Qy 454 CSDYEDTEE--RPRPQFAASAPMTAPNPITGEDEPYFPERSRARMLAGSVVIVVMVAVV 511
 Db 395 LVDFEEEQQLQLRPEYEA KCTDKKNPVTQEMEPYLPSSKAVRFCFSGATVLFWISLI 454

Qy 512 VMCLVSIILYRA-IMAI VVSRS GNTLLAAWASRI-----ASLTGSVNLVFIILISKIY 564
 Db 455 IASIIAIIYRLWVYAAFASIMENNLTLPEVRNLLTPQLATSVTASVLNFIITIMILNFLY 514

Qy 565 VSLAHLVLRWEMHRTQTKFEDAFTLKVFIFQFVNFYSSPVYIAFFKGRFVGYPGNYHTLF 624
 Db 515 ERVIAIWITDMEIPRTHLEYENRLTMKMLFQFVNYSSCFYVAFKGFVGYPADYTYLF 574

Qy 625 GV-RNEECAAGGCLIELAQELLVIMVGKQVINNMQEVLI PKLKGWQKFRLRSKRRKAGA 683
 Db 575 GKWRNEECDPAGCLIELTQTITIVMAGKQIWGNIQEA FVPWTWNW-----LKRRKARN 627

Qy 684 SAGASQGPWEDDYELVPE--GLFDEYLEMVLQFGFVTIFVAACPLAPL FALLNNWVEIR 741
 Db 628 HPENLYSRWEQDGLQTFGGGLGFYEYLEMVVQFGFITLFVASFPLAPLLALLNNILEIR 687

Qy 742 LDARKFVCEYRRPVAERAQDIGIWFHILAGLTHLAVISNAFL LAFSSDFLPRA--YYRWT 799
 Db 688 VDSWKLTTQFKRPVAAKAHSIGVWQEILNGIAILSVVTNAFIVAFTSDMI PRLVYYYAYT 747

Qy 800 RAHD--LRGFLNFTLARAPSSFAAAHNR-----TCRYRAFRDDDGH---YSQ 841
 Db 748 QDKDMPMSGYISSSL----SIFNVTDFKEQSMPTKNDMNVLSCRYRDYRYP PGHGKEYQV 803

Qy 842 T--YWNLLAIRLAFVIVFEHV VFSVGRLLDLLVPDIPESVEIKVKREYYLAKQALAENEV 899
 Db 804 TMQYWHILAAKMAFIIIMEHV VFLVKFFVAWLIPDIPSEVKARVKREKFLTQKILHEYEL 863

RESULT 14

Q9VTS0_DROME

ID Q9VTS0_DROME PRELIMINARY; PRT; 1219 AA.
 AC Q9VTS0;
 DT 01-MAY-2000, integrated into UniProtKB/TreMBL.
 DT 01-MAY-2000, sequence version 1.
 DT 07-FEB-2006, entry version 25.
 DE CG6938-PA.
 GN ORFNames=CG6938, Dmel_CG6938;
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
 RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=22426065; PubMed=12537568;
 RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
 RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
 RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,
 RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
 RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
 RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
 RT "Finishing a whole-genome shotgun: release 3 of the *Drosophila*
 RT *melanogaster* euchromatic genome sequence.";
 RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=22426070; PubMed=12537573;
 RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J.W., Svirskas R.,
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
 RA Ashburner M., Celniker S.E.;
 RT "The transposable elements of the *Drosophila melanogaster* euchromatin:
 RT a genomics perspective.";
 RL Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).
 RN [4]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=22426069; PubMed=12537572;
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
 RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.;
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a
 RT systematic review.";
 RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
 RN [5]
 RP NUCLEOTIDE SEQUENCE.
 RG Berkeley *Drosophila* Genome Project;
 RA Celniker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,
 RA Hoskins R., Stapleton M., Pacleb J., Park S., Svirskas R., Smith E.,
 RA Yu C., Rubin G.;

```

RT      "Drosophila melanogaster release 4 sequence.";
RL      Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN      [6]
RP      NUCLEOTIDE SEQUENCE.
RG      FlyBase;
RL      Submitted (JAN-2006) to the EMBL/GenBank/DDBJ databases.
CC      -----
CC      Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC      Distributed under the Creative Commons Attribution-NoDerivs License.
CC      -----
DR      EMBL; AE003543; AAF49976.1; -; Genomic_DNA.
DR      FlyBase; FBgn0036235; CG6938.
DR      InterPro; IPR007632; DUF590.
DR      InterPro; IPR003006; Ig_MHC.
DR      InterPro; IPR001563; Peptidase_S10.
DR      Pfam; PF04547; DUF590; 1.
DR      PROSITE; PS00560; CARBOXYPEPT_SER_HIS; UNKNOWN_1.
DR      PROSITE; PS00290; IG_MHC; UNKNOWN_1.
SO      SEQUENCE 1219 AA; 139592 MW; F5ABEB2726ED82A6 CRC64;

```

Qy	35	AERWAMTSETSSGSHCARSRML-----RRRAQEEDSTVLIDVSPPEAEKRGSY	82
Db	249	ADRVNQSYEVMESSH---SNVLPDQFGYRQLIPTERKASDTASSV-----SGSY	294
Qy	83	GSTAHASEP---GGQQAACRAGSPAKP-----RIADFVLVW-EEDLKLDRQ	125
Db	295	YGSRKASKSNSLGGESGDERRVSKQDREGLDPESLMFRDGRKQVDMVLAWEEEDLGMTE	354
Qy	126	QDSAARDRTDMHRTWRETFLDNLRAAGLCVDQOD-VQDGNTHVHYALLSASWAVLCYYAE	184
Db	355	AEAKRRDN-----RRSFMENLIKEGLEVELEDKQSFNEKTFFLKIHLPWRLLETRLAE	407
Qy	185	DLRLKLP-----LQELPNQASNWSAGLLAWLGIPNVLLLEVVPDVP	225
Db	408	VMNLKLPVKRFITISVKPSWDEENVVLRNMQYWKDVWQR-LTKKIQLDQTLLE---GET	462
Qy	226	EYYSRFRVKNLPRFLGSDNQDTFFTSTKRHQILFEILAKTPYGHEKKNLLGIHQLLAEG	285
Db	463	TFKAATANGNPEEQFIVKD-RATAFTSAQRSLMVMQVLIRTPFDESRS--GIRRLMNDG	519
Qy	286	VLSAAPFLHDGPFKTPPEGFPQAPRLN-QRQVLFQHWARWGKWNKYQPLDHYVRRYFGEKVA	344
Db	520	TYLGCFFLHEGRY---DRPHSSGISLDRRVLYQTWAHPSQWYKKQPLCLVRKYFGDKIA	575
Qy	345	LYFAWLGFYTGWLLPAAVVGTLVFLVGCFLVFS--IPTQELCG--SKDSFEMCPLC-LD	399
Db	576	LYFCWLGFYTEMLVYPVAVGTLCFIYGATLESEDNTPSKEICNEYGTGNITLCLCDKA	635
Qy	400	CPFWLLSSACALAQAGRLFDHGGTVFFSLFMALWAVLLELYWKRKSATLAYRWDCSDYED	459
Db	636	CSYQRLSESCFLSRLTYLFDNPSTVFFAIFMSFWATTFLLELWKRKQSVLVWEWDLHNV-D	694
Qy	460	TEERPRPQFAASAPMTAPNPITGEDEPYFPERSRARRMLAGSVVIVVMVAVVMCLVSII	519
Db	695	MDEENRPEFETNATTFRMNPVTREKEPYMSTWNRSIRFVITGSAVLFMISVLSAVLGTI	754
Qy	520	LYRAIMAIVVSRSNTLLAAWASRIASLTGSVVNLVFIILSKIYVSLAHVLTREWHR	579
Db	755	LYRITLVSVIYGGGGFFVKEHAKLFTSVTAALINLVVIMILTRIYHRMAIKLTNLENPT	814
Qy	580	QTKFEDAFTLKVFIFQFVNFYSSPVYIAFFKGRFVGYPGNYHT----LFGVRNEECAAGG	635
Db	815	HTEYEDSYTFKIFFFEFNMFYSSLIYIAFFKGRFFDYPGDDQARKSEFFRLKNDICDPAG	874
Qy	636	CLIELAQELLVIMVGKQVINNMQEVLPKLGKWWQKFRLRSSKKRKAGASAGASQGPWEDD	695
Db	875	CLSELCIQLAIMVGKQCWNFMMEYLPFKFWNWWR----QRKHQATKDESHLMMAEQD	930
Qy	696	YELV-PCE-GLFDEYLEMVLQFGFVTIFVAACPLAPLALLNNWVEIRLDARKFVCEYRR	753
Db	931	YHMQDPGRRLALFDEYLEMILQYGFVTLFVAAPLAPLALLNNVAEIRLDAYKMVTQARR	990

Qy 754 PVAERAQDIGIWFHILAGLTHLAVISNAFLAFSSDFLPRAYR--WTRAHDLRGFLNFT 811
 Db 991 PLAERVEDIGAWYGILRIITYTAVVSNAFVIAYTSDFIPRMVYKFVYSETHTLAGYIEHS 1050

Qy 812 LA-----RAPSSFAAAHNRTCRYAFRDDGHHY-----SQTYWNLALAIRLAFVIVF 857
 Db 1051 LSIFNTSDYKEEWGASVSEKDPDTCQYRGYRNGPKDYEPYGLSPHYWHVFAARLAFVVVF 1110

Qy 858 EHVVSFVGRLLDLLVPDIPESVEIKVKREYYLAKQALAENEVLFGTNGTKDEQPKGSELS 917
 Db 1111 EHVVFVITGIMQFIIPDVPSEVKTQMREQLLAKEAKYQ-----HGKRAQGDSQDIM 1163

Qy 918 S 918
 Db 1164 S 1164

RESULT 15

Q2MOY5_DROPS

ID Q2MOY5_DROPS PRELIMINARY; PRT; 1235 AA.
 AC Q2MOY5;
 DT 21-FEB-2006, integrated into UniProtKB/TrEMBL.
 DT 21-FEB-2006, sequence version 1.
 DT 21-FEB-2006, entry version 1.
 DE GA19969-PA (Fragment).
 GN Name=Dpse\GA19969; ORFNames=Dpse_GA19969;
 OS Drosophila pseudoobscura (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7237;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=MV2-25;
 RX PubMed=15632085; DOI=10.1101/gr.3059305;
 RA Richards S., Liu Y., Bettencourt B.R., Hradecky P., Letovsky S.,
 RA Nielsen R., Thornton K., Hubisz M.J., Chen R., Meisel R.P.,
 RA Couronne O., Hua S., Smith M.A., Zhang P., Liu J., Bussemaker H.J.,
 RA van Batenburg M.F., Howells S.L., Scherer S.E., Sodergren E.,
 RA Matthews B.B., Crosby M.A., Schroeder A.J., Ortiz-Barrientos D.,
 RA Rives C.M., Metzker M.L., Muzny D.M., Scott G., Steffen D.,
 RA Wheeler D.A., Worley K.C., Havlak P., Durbin K.J., Egan A., Gill R.,
 RA Hume J., Morgan M.B., Miner G., Hamilton C., Huang Y., Waldron L.,
 RA Verduzco D., Clerc-Blankenburg K.P., Dubchak I., Noor M.A.F.,
 RA Anderson W., White K.P., Clark A.G., Schaeffer S.W., Gelbart W.,
 RA Weinstock G.M., Gibbs R.A.;
 RT "Comparative genome sequencing of Drosophila pseudoobscura:
 RT chromosomal, gene, and cis-element evolution."
 RL Genome Res. 15:1-18(2005).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=MV2-25;
 RG FlyBase;
 RL Submitted (JUN-2005) to the EMBL/GenBank/DDBJ databases.
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=MV2-25;
 RG Human Genome Sequencing Center;
 RA Richards S., Liu Y., Bettencourt B.R., Hradesky P., Letovsky S.,
 RA Chen R., Smith M.A., Howells S.L., Scherer S.E., Sodergren E.,
 RA Rives C.M., Metzker M.L., Munzy D.M., Wheeler D.A., Worley K.C.,
 RA Havlak P., Durbin K.J., Egan A., Gill R., Hume J., Morgan M.B.,
 RA Huang Y., Waldron L., Verduzco D., Blankenburg K.P., Adams C.,
 RA Allen C., Allen H., Anyalebechi V., Asomugha C., Bellard T.,
 RA Bhuchar V., Biswalo K., Blair J., Blomstrom D., Burrell K.,
 RA Calderon E., Cardenas V., Carter K., Cavazos I., Ceasar H., Chacko J.,
 RA Chavez D., Chu J., Cockrell R., Cox C., Coyle M., Davila M., Davis C.,
 RA Davy-Carroll L., De A., Delgado O., Denson S., Deramo C., Dinh H.,
 RA Eaves K., Escotto M., Eugene C., Falls T., Fernandez S., Flagg N.,
 RA Forbes L., Garner T., Garza M., Ghose S., Grady M., Hamilton C.,
 RA Hernandez J., Hines S., Hogues M., Hollins B., Idlebird D., Imo K.,
 RA Jimenez A., Johnson B., Jolivet A., Kelly S., King L., Kisamo H.,
 RA Kovar C., Lebow H., lee K., LeGall F., Lewis L., Li Z., London P.,
 RA Lopez J., Lozado R., Malloy K., Martinez E., Mercadao C., Miner G.,
 RA Minja E., Moore S., Nanavati A., Ngo R., Nguyen N., Nwaokemele O.,
 RA Okwuonu G., Parks K., Pasternak S., Patel B., Paul H., Payne C.,

<http://es/ScoreAccessWeb/GetItem.action?AppId=10552515&seqId=775630&ItemName...> 11/17/2006

Search completed: October 27, 2006, 20:28:46
Job time : 327 secs

<http://es/ScoreAccessWeb/GetItem.action?AppId=10552515&seqId=775630&ItemName...> 11/17/2006

SCORE Search Results Details for Application 10552515 and Search Result us-10-552-515-1.rag.

[Score Home Page](#)
[Retrieve Application List](#)
[SCORE System Overview](#)
[SCORE FAQ](#)
[Comments / Suggestions](#)

This page gives you Search Results detail for the Application 10552515 and Search Result us-10-552-515-1.rag.

[start](#)

[Go Back to previous page](#)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2006, 20:18:31 ; Search time 205 Seconds
(without alignments)
2080.892 Million cell updates/sec

Title: US-10-552-515-1
Perfect score: 4950
Sequence: 1 MRMAATAWAGLQGPPPLPTLC.....SELSSHWTPFTVPKASQLQQ 933

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_8:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*
10: geneseqp2006s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	4950	100.0	933	8 ADT77664	Adt77664 Splice va
2	4531.5	91.5	885	9 AEB13426	Aeb13426 Human pro
3	4364.5	88.2	843	9 AEB13424	Aeb13424 Human pro
4	3736	75.5	898	4 ABG15488	Abg15488 Novel hum
5	1531.5	30.9	920	7 ADB64420	Adb64420 Human pro
6	1511.5	30.5	920	6 ABP58666	Abp58666 Human dih
7	1504	30.4	981	8 ADK52114	Adk52114 Human ato

8	1488	30.1	960	10	AEG11142	Aeg11142 Human tra
9	1479.5	29.9	840	10	AEG11146	Aeg11146 Human tra
10	1464	29.6	1003	7	ADG48280	Adg48280 Human ret
11	1445	29.2	1219	4	ABB62812	Abb62812 Drosophil
12	1402.5	28.3	910	7	ADC42854	Adc42854 REMAP pro
13	1369.5	27.7	1075	4	ABB65993	Abb65993 Drosophil
14	1367.5	27.6	712	10	AEG11145	Aeg11145 Human tra
15	1199.5	24.2	1058	4	ABB65022	Abb65022 Drosophil
16	1154	23.3	596	7	ADB64387	Adb64387 Human pro
17	1061.5	21.4	594	4	AAB92637	Aab92637 Human pro
18	1061.5	21.4	594	5	ABP43811	Abp43811 FLJ10261
19	1061.5	21.4	594	8	ADJ75429	Adj75429 Marker ge
20	1061.5	21.4	594	8	ADN04848	Adn04848 Antipsori
21	1061.5	21.4	594	10	AEG11143	Aeg11143 Human FLJ
22	1037.5	21.0	782	7	ADT95905	Adt95905 Colon can
23	1037.5	21.0	782	7	ADX42387	Adx42387 Human col
24	1037.5	21.0	782	8	ADQ96288	Adq96288 T cell ac
25	1037.5	21.0	782	8	ADQ96104	Adq96104 T cell ac
26	912.5	18.4	475	7	ADB64962	Adb64962 Human pro
27	905	18.3	642	7	ADM05798	Adm05798 Human pro
28	905	18.3	642	9	AEC88728	Aec88728 Human cDN
29	905	18.3	642	10	AEG11144	Aeg11144 Human FLJ
30	819.5	16.6	443	5	ABP41785	Abp41785 Human ova
31	817.5	16.5	179	7	AAO29613	Aao29613 Human Nov
32	784.5	15.8	390	5	ABB90382	Abb90382 Human pol
33	735	14.8	139	5	AAE24066	Aae24066 Human pro
34	722.5	14.6	360	4	AAM40391	Aam40391 Human pol
35	711.5	14.4	346	8	ADP29628	Adp29628 Human sec
36	695.5	14.1	608	8	ADQ96298	Adq96298 T cell ac
37	695.5	14.1	608	8	ADQ96286	Adq96286 T cell ac
38	684.5	13.8	483	7	ADM05305	Adm05305 Human pro
39	684.5	13.8	483	8	ADQ96290	Adq96290 T cell ac
40	684.5	13.8	483	9	AEC88235	Aec88235 Human cDN
41	656.5	13.3	314	4	AAM42177	Aam42177 Human pol
42	612	12.4	339	4	AAB94837	Aab94837 Human pro
43	601	12.1	594	8	ADQ67527	Adq67527 Novel hum
44	594.5	12.0	589	4	AAB92752	Aab92752 Human pro
45	594.5	12.0	589	5	ABB97370	Abb97370 Novel hum

ALIGNMENTS

RESULT 1

ADT77664

ID ADT77664 standard; protein; 933 AA.

XX

AC ADT77664;

XX

DT 13-JAN-2005 (first entry)

XX

DE Splice variant-novel gene expressed in prostate (SV-NGEP) polypeptide.

XX

KW Splice variant-novel gene expressed in prostate; SV-NGEP; human;
prostate cancer; cytostatic; gene therapy; immunotherapy.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT Domain 1. .345

FT /label = Cytoplasmic

FT Region 157. .933

FT /note= "An immunogenic fragment comprising 8 consecutive
amino acids that specifically binds to an antibody that
specifically binds to a polypeptide comprising amino
acids 157-933 is referred to in Claim 1"

FT Region 170. .178

FT /note= "Epitope, predicted to bind HLA2-01"

FT Region 215. .223

FT /note= "Epitope, predicted to bind HLA2-01"

FT Region 258. .266

FT /note= "Epitope, predicted to bind HLA2-01"

FT Domain 346. .368

FT /label = Transmembrane

FT Domain 369. .421

FT /label = External

FT /note= "Cell surface"
 FT Region 403. .411
 FT /note= "Epitope, predicted to bind HLA2-01"
 FT Domain 422. .441
 FT /label = Transmembrane
 FT Region 427. .435
 FT /note= "Epitope, predicted to bind HLA2-01"
 FT Domain 442. .501
 FT /label = Cytoplasmic
 FT Domain 502. .524
 FT /label = Transmembrane
 FT Domain 525. .543
 FT /label = External
 FT /note= "Cell surface"
 FT Domain 544. .566
 FT /label = Transmembrane
 FT Region 557. .565
 FT /note= "Epitope, predicted to bind HLA2-01"
 FT Region 562. .570
 FT /note= "Epitope, predicted to bind HLA2-01"
 FT Domain 567. .586
 FT /label = Cytoplasmic
 FT Domain 587. .609
 FT /label = Transmembrane
 FT Domain 610. .714
 FT /label = External
 FT /note= "Cell surface"
 FT Domain 715. .737
 FT /label = Transmembrane
 FT Domain 738. .761
 FT /label = Cytoplasmic
 FT Domain 762. .784
 FT /label = Transmembrane
 FT Domain 785. .933
 FT /label = External
 FT /note= "Cell surface"
 FT Region 846. .854
 FT /note= "Epitope, predicted to bind HLA2-01"
 XX
 PN WO2004092213-A1.
 XX
 PD 28-OCT-2004.
 XX
 PF 05-APR-2004; 2004WO-US010588.
 XX
 PR 08-APR-2003; 2003US-0461399P.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Pastan I, Bera TK, Lee B;
 XX
 DR WPI; 2004-758338/74.
 DR N-PSDB; ADT77665.
 XX
 PT New Splice Variant-Novel Gene Expressed in Prostate polypeptide or
 PT encoding nucleic acid molecule for diagnosing, preventing or treating
 PT cancer, especially prostate cancer.
 XX
 PS Claim 1; SEQ ID NO 1; 88pp; English.
 XX
 CC The present sequence is the protein sequence of splice variant-novel gene
 CC expressed in prostate (SV-NGEP). SV-NGEP is identical to NGEP from amino
 CC acid 1-157, diverging from amino acid 158. Expression analysis in 76
 CC normal and foetal tissues showed SV-NGEP to be strongly expressed only in
 CC a prostate sample. Claimed methods for detecting prostate cancer in a
 CC subject comprise: contacting the sample with an antibody that
 CC specifically binds a SV-NGEP polypeptide and detecting the formation of
 CC an immune complex; or detecting an increase in expression of SV-NGEP
 CC polypeptide or mRNA. Antibodies to an SV-NGEP polypeptide can be used to
 CC detect metastatic prostate cancer cells at locations other than the
 CC prostate. A claimed method for producing an immune response against a
 CC cell expressing SV-NGEP, for example in a subject with prostate cancer,
 CC comprises administering the polypeptide, or a polynucleotide encoding it,
 CC to produce an immune response that decreases growth of the prostate
 CC cancer. A claimed method for inhibiting the growth of a malignant cell
 CC that expresses SV-NGEP comprises culturing cytotoxic T lymphocytes (CTLs)

CC with SV-NGEP to produce activated CTLs that recognise an NGEF expressing
 CC cell, and contacting the malignant cell with the activated CTLs.
 CC Alternatively, growth of a malignant cell is inhibited by contact with an
 CC antibody that specifically binds an SV-NGEP polypeptide, where the
 CC antibody is linked to an effector molecule (chemotherapeutic agent or
 CC toxin) that inhibits growth of the malignant cell. This may be performed
 CC in vivo. Kits for detecting an SV-NGEP polypeptide or polynucleotide in a
 CC sample are also claimed.
 XX
 SQ Sequence 933 AA;

Query Match 100.0%; Score 4950; DB 8; Length 933;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 933; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy      1 MRMAATAWAGLQGPPLPTLCPAVRTGLYCRDQAHAEWAMTSETSSGSHCARSMLRRRA 60
      |||
Db      1 MRMAATAWAGLQGPPLPTLCPAVRTGLYCRDQAHAEWAMTSETSSGSHCARSMLRRRA 60

Qy     61 QEEDSTVLIDVSPPEAEKRGSGYSTAHASEPGGQAAACRAGSPAKPRIADFVLVWEEDL 120
      |||
Db     61 QEEDSTVLIDVSPPEAEKRGSGYSTAHASEPGGQAAACRAGSPAKPRIADFVLVWEEDL 120

Qy    121 KLDRQQDSAADRTDMHRTWRETFLDNLRAAGLCVDQDQDGNNTVHYALLSASWAVLC 180
      |||
Db    121 KLDRQQDSAADRTDMHRTWRETFLDNLRAAGLCVDQDQDGNNTVHYALLSASWAVLC 180

Qy    181 YYAEDLRLKLPLQELPNQASNWSAGLLAWLGIPNVLLEVVPDVPPEYYSCRFRVKNLPRF 240
      |||
Db    181 YYAEDLRLKLPLQELPNQASNWSAGLLAWLGIPNVLLEVVPDVPPEYYSCRFRVKNLPRF 240

Qy    241 LGSDNQDTFFTSTKRHQILFEILAKTPYGHEKKNLLGIHQLLAEGVLSAAFPLHDGPFKT 300
      |||
Db    241 LGSDNQDTFFTSTKRHQILFEILAKTPYGHEKKNLLGIHQLLAEGVLSAAFPLHDGPFKT 300

Qy    301 PPEGQPAPRLNQRQVLFQHWARWGKWNKYQPLDHVRRYFGEKVALYFAWLGFYTGWLLPA 360
      |||
Db    301 PPEGQPAPRLNQRQVLFQHWARWGKWNKYQPLDHVRRYFGEKVALYFAWLGFYTGWLLPA 360

Qy    361 AVVGTLVFLVGCFLVFSDIPTQELCGSKDSFEMCPLCLDCPFWLLSSACALAQAGRLFDH 420
      |||
Db    361 AVVGTLVFLVGCFLVFSDIPTQELCGSKDSFEMCPLCLDCPFWLLSSACALAQAGRLFDH 420

Qy    421 GGTVFFSLFMALWAVLLLEYWKRKSATLAYRWDCSDYEDTEERPRPQFAASAPMTAPNPI 480
      |||
Db    421 GGTVFFSLFMALWAVLLLEYWKRKSATLAYRWDCSDYEDTEERPRPQFAASAPMTAPNPI 480

Qy    481 TGEDEPYFPERSRARRMLAGSVVIVVMVAVVMCLVSIILYRAIMAIIVSRSGNTLLAAW 540
      |||
Db    481 TGEDEPYFPERSRARRMLAGSVVIVVMVAVVMCLVSIILYRAIMAIIVSRSGNTLLAAW 540

Qy    541 ASRIASLTGSVNVLVFILILSKIYVSLAHVLTWRWEMHRTQTKFEDAFTLKVFIFQFVNFY 600
      |||
Db    541 ASRIASLTGSVNVLVFILILSKIYVSLAHVLTWRWEMHRTQTKFEDAFTLKVFIFQFVNFY 600

Qy    601 SSPVYIAFFKGRFVGYPGNYHTLFGVRNEECAAGGCLIELAQELLVIMVGKQVINNMQEV 660
      |||
Db    601 SSPVYIAFFKGRFVGYPGNYHTLFGVRNEECAAGGCLIELAQELLVIMVGKQVINNMQEV 660

Qy    661 LIPKLKGWQKFRRLRSKRRKAGASAGSQGPWEDDYELVPCEGLFDEYLEMVLQFGFVTI 720
      |||
Db    661 LIPKLKGWQKFRRLRSKRRKAGASAGSQGPWEDDYELVPCEGLFDEYLEMVLQFGFVTI 720

Qy    721 FVAACPLAPLFA LLNNWVEIRLDARKFVCEYRRPVAERAQDIGIWFHILAGLTHLAVISN 780
      |||
Db    721 FVAACPLAPLFA LLNNWVEIRLDARKFVCEYRRPVAERAQDIGIWFHILAGLTHLAVISN 780

Qy    781 AFLAFSSDFLPRAYYRWTRAHDLRGFLNFTLARAPSSFAAAHNRTCRYRAFRDDDGHS 840
      |||
Db    781 AFLAFSSDFLPRAYYRWTRAHDLRGFLNFTLARAPSSFAAAHNRTCRYRAFRDDDGHS 840

Qy    841 QTYWNLALIRLAFVIVFEHVVSFVGRLLDLVDPDIPESVEIKVKREYYLAKQALAENEVL 900
      |||
Db    841 QTYWNLALIRLAFVIVFEHVVSFVGRLLDLVDPDIPESVEIKVKREYYLAKQALAENEVL 900

Qy    901 FGTNGTKDEQPKGSELSSHWPFTVPKASQLQQ 933

```

|||||
Db 901 FGNGTKDEQPKGSELSSHWTPTVTPKASQLQQ 933

RESULT 2

AEB13426

ID AEB13426 standard; protein; 885 AA.

XX

AC AEB13426;

XX

DT 22-SEP-2005 (first entry)

XX

DE Human prostate specific polypeptide #2.

XX

KW Screening; diagnosis; drug delivery; prostate specific polypeptide;

KW cancer; prostate tumor; cytostatic; neoplasm.

XX

OS Homo sapiens.

XX

PN WO2005062788-A2.

XX

PD 14-JUL-2005.

XX

PF 16-DEC-2004; 2004WO-US042406.

XX

PR 22-DEC-2003; 2003US-0531809P.

XX

PA (AVAL-) AVALON PHARM INC.

XX

PI Weigle B, Ebner R;

XX

DR WPI; 2005-497793/50.

DR

N-PSDB; AEB13425.

XX

PT Novel isolated prostate specific polypeptide, useful for treating cancer,

PT and identifying agent that modulates activity of cancer related gene.

XX

PS Claim 12; SEQ ID NO 5; 59pp; English.

XX

CC The invention relates to an isolated prostate specific polypeptide
CC comprising one or more immunogenic fragments. The invention also relates
CC to a method of identifying an agent that modulates the activity of a
CC cancer related gene involving contacting a compound with a cell
CC containing a gene under conditions promoting the expression of the gene,
CC detecting a difference in expression of the gene relative to when the
CC compound is not present and identifying an agent that modulates the
CC activity of a cancer related gene, a method of identifying an anti-
CC neoplastic agent involving contacting a cell exhibiting neoplastic
CC activity with a compound first identified as a cancer related gene
CC modulator using and determining a decrease in neoplastic activity after
CC contacting, when compared to when the contacting does not occur, or
CC administering an agent first identified to an animal exhibiting a cancer
CC condition and detecting a decrease in cancerous condition, a method of
CC determining the cancerous status of a cell involving determining an
CC increase in the level of expression in a cell of a gene where an elevated
CC expression relative to a known non-cancerous cell indicates a cancerous
CC state or potentially cancerous state, an antibody that reacts with a
CC prostate specific polypeptide, an immunoconjugate comprising the antibody
CC and a cytotoxic agent, a method of treating cancer involving contacting a
CC cancerous cell in vivo with an agent having activity against a prostate
CC specific polypeptide and an immunogenic composition the prostate specific
CC polypeptide. The prostate specific polypeptide is useful for identifying
CC an agent that modulates the activity of a cancer related gene. The
CC immunogenic composition is useful for treating cancer, preferably
CC prostate cancer in an animal, e.g. human, which involves administering
CC the immunogenic composition that is sufficient to elicit the production
CC of cytotoxic T lymphocytes specific for the prostate specific
CC polypeptide. The invention is useful for identifying anti-neoplastic
CC agents. This sequence represents a human prostate specific polypeptide of
CC the invention.

XX

SQ Sequence 885 AA;

Query Match 91.5%; Score 4531.5; DB 9; Length 885;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 855; Conservative 0; Mismatches 0; Indels 3; Gaps 2;

Qy 1 MRMAATAWAGLQGPPLPTLCPAVRTGLYCRDQQAHAERWAMTSETSSGSHCARSRMLRRRA 60
 Db 5 MRMAATAWAGLQGPPLPTLCPAVRTGLYCRDQQAHAERWAMTSETSSGSHCA--RMLRRRA 62

Qy 61 QEEDSTVLIDVSPPEAEKRGSYGSTAHASEPGGQQAACRAGSPAKPRIADFLVWEEEDL 120
 Db 63 QEEDSTVLIDVSPPEAEKRGSYGSTAHASEPGGQQAACRAGSPAKPRI-DFVLVWEEEDL 121

Qy 121 KLD RQDQSAARDRTDMHRTWRETFLDNLRAAGLCVDQDQDGNNTTVHYALLSASWAVLC 180
 Db 122 KLD RQDQSAARDRTDMHRTWRETFLDNLRAAGLCVDQDQDGNNTTVHYALLSASWAVLC 181

Qy 181 YYAEDLRLKLPLQELPNQASNWSAGLLAWLGIPNVLLEVVPDVPPEYSCRFRVKNLPRF 240
 Db 182 YYAEDLRLKLPLQELPNQASNWSAGLLAWLGIPNVLLEVVPDVPPEYSCRFRVKNLPRF 241

Qy 241 LGS DNQDTFTTSTKRHQILFEILAKTPYGHEKKNLLGIHQLLAEGVLSAAFPLHDGPFT 300
 Db 242 LGS DNQDTFTTSTKRHQILFEILAKTPYGHEKKNLLGIHQLLAEGVLSAAFPLHDGPFT 301

Qy 301 PPEGPQAPRLNQRQVLFQHWARWGKWNKYQPLDHVRRYFGEKVALYFAWLGFYTGWLLPA 360
 Db 302 PPEGPQAPRLNQRQVLFQHWARWGKWNKYQPLDHVRRYFGEKVALYFAWLGFYTGWLLPA 361

Qy 361 AVVGTLVFLVGCFLVFSDIPTQELCGSKDSFEMCPLCLDCPFWLLSSACALAQAGRLFDH 420
 Db 362 AVVGTLVFLVGCFLVFSDIPTQELCGSKDSFEMCPLCLDCPFWLLSSACALAQAGRLFDH 421

Qy 421 GGTVFVFSLFMALWAVLLLEYWKRKSATLAYRWDCSDYEDTEERPRPQFAASAPMTAPNPI 480
 Db 422 GGTVFVFSLFMALWAVLLLEYWKRKSATLAYRWDCSDYEDTEERPRPQFAASAPMTAPNPI 481

Qy 481 TGEDEPYFPERSRARRMLAGSVVIVVMVAVVMCLVSIILYRAIMAIVVSRSGNTLLAAW 540
 Db 482 TGEDEPYFPERSRARRMLAGSVVIVVMVAVVMCLVSIILYRAIMAIVVSRSGNTLLAAW 541

Qy 541 ASRIASLTGSVNVLVFILILSKIYVSLAHVLT RWEMHRTQTKFEDAFTLKVFIFQFVN FY 600
 Db 542 ASRIASLTGSVNVLVFILILSKIYVSLAHVLT RWEMHRTQTKFEDAFTLKVFIFQFVN FY 601

Qy 601 SSPVYIAFFKGRFVGYPGNYHTLFGVRNEECAAGGCLIELAQELLVIMVGKQVINNMQEV 660
 Db 602 SSPVYIAFFKGRFVGYPGNYHTLFGVRNEECAAGGCLIELAQELLVIMVGKQVINNMQEV 661

Qy 661 LIPKLGWQKFR LRSKKRKAGASAGASQGPWEDDYELVPCEGLFDEYLEMVLQFGFVTI 720
 Db 662 LIPKLGWQKFR LRSKKRKAGASAGASQGPWEDDYELVPCEGLFDEYLEMVLQFGFVTI 721

Qy 721 FVAACPLAPL FALLNNWVEIRLDARKEVCEYRRPVAERAQDIGIWFHILAGLTHLAVISN 780
 Db 722 FVAACPLAPL FALLNNWVEIRLDARKEVCEYRRPVAERAQDIGIWFHILAGLTHLAVISN 781

Qy 781 AFL LAFSSDFLP RAYYRWTRAHDLRGFLNFTLARAPSSFAAAHNRTC RYAFRDDDGHYS 840
 Db 782 AFL LAFSSDFLP RAYYRWTRAHDLRGFLNFTLARAPSSFAAAHNRTC RYAFRDDDGHYS 841

Qy 841 QTYWNLLAIRLAFVIVFE 858
 Db 842 QTYWNLLAIRLAFVIVFE 859

RESULT 3

AEB13424

ID AEB13424 standard; protein; 843 AA.

XX

AC AEB13424;

XX

DT 22-SEP-2005 (first entry)

XX

DE Human prostate specific polypeptide #1.

XX

KW Screening; diagnosis; drug delivery; prostate specific polypeptide;
 KW cancer; prostate tumor; cytostatic; neoplasm.

XX

OS Homo sapiens.

XX
 PN WO2005062788-A2.
 XX
 PD 14-JUL-2005.
 XX
 PF 16-DEC-2004; 2004WO-US042406.
 XX
 PR 22-DEC-2003; 2003US-0531809P.
 XX
 PA (AVAL-) AVALON PHARM INC.
 XX
 PI Weigle B, Ebner R;
 XX
 DR WPI; 2005-497793/50.
 DR N-PSDB; AEB13423.
 XX
 PT Novel isolated prostate specific polypeptide, useful for treating cancer,
 PT and identifying agent that modulates activity of cancer related gene.
 XX
 PS Claim 12; SEQ ID NO 3; 59pp; English.
 XX
 CC The invention relates to an isolated prostate specific polypeptide
 CC comprising one or more immunogenic fragments. The invention also relates
 CC to a method of identifying an agent that modulates the activity of a
 CC cancer related gene involving contacting a compound with a cell
 CC containing a gene under conditions promoting the expression of the gene,
 CC detecting a difference in expression of the gene relative to when the
 CC compound is not present and identifying an agent that modulates the
 CC activity of a cancer related gene, a method of identifying an anti-
 CC neoplastic agent involving contacting a cell exhibiting neoplastic
 CC activity with a compound first identified as a cancer related gene
 CC modulator using and determining a decrease in neoplastic activity after
 CC contacting, when compared to when the contacting does not occur, or
 CC administering an agent first identified to an animal exhibiting a cancer
 CC condition and detecting a decrease in cancerous condition, a method of
 CC determining the cancerous status of a cell involving determining an
 CC increase in the level of expression in a cell of a gene where an elevated
 CC expression relative to a known non-cancerous cell indicates a cancerous
 CC state or potentially cancerous state, an antibody that reacts with a
 CC prostate specific polypeptide, an immunoconjugate comprising the antibody
 CC and a cytotoxic agent, a method of treating cancer involving contacting a
 CC cancerous cell in vivo with an agent having activity against a prostate
 CC specific polypeptide and an immunogenic composition the prostate specific
 CC polypeptide. The prostate specific polypeptide is useful for identifying
 CC an agent that modulates the activity of a cancer related gene. The
 CC immunogenic composition is useful for treating cancer, preferably
 CC prostate cancer in an animal, e.g. human, which involves administering
 CC the immunogenic composition that is sufficient to elicit the production
 CC of cytotoxic T lymphocytes specific for the prostate specific
 CC polypeptide. The invention is useful for identifying anti-neoplastic
 CC agents. This sequence represents a human prostate specific polypeptide of
 CC the invention.
 XX
 SQ Sequence 843 AA;

Query Match 88.2%; Score 4364.5; DB 9; Length 843;
 Best Local Similarity 99.6%; Pred. No. 0;
 Matches 824; Conservative 0; Mismatches 0; Indels 3; Gaps 2;

Qy 1 MRMAATAWAGLQGPPLPTLCPAVRTGLYCRDQAHAEWAMTSETSSGSHCARSRMLRRRA 60
 |||||
 Db 5 MRMAATAWAGLQGPPLPTLCPAVRTGLYCRDQAHAEWAMTSETSSGSHCA--RMLRRRA 62
 Qy 61 QEEDSTVLIDVSPPEAEKRGSYGSTAHASEPGGQQAACRAGSPAKPRIADFLVWEEEDL 120
 |||||
 Db 63 QEEDSTVLIDVSPPEAEKRGSYGSTAHASEPGGQQAACRAGSPAKPRI-DFVLVWEEEDL 121
 Qy 121 KLDROQDSAARDRTDMHRTWRETFLDNLRAAGLCVDQDQDGNNTTVHYALLSASWAVLC 180
 |||||
 Db 122 KLDROQDSAARDRTDMHRTWRETFLDNLRAAGLCVDQDQDGNNTTVHYALLSASWAVLC 181
 Qy 181 YYAEDLRLKLPLQELPNQASNWSAGLLAWLGIPNVLLEVVPDVPPEYYSRFRVKNLPRF 240
 |||||
 Db 182 YYAEDLRLKLPLQELPNQASNWSAGLLAWLGIPNVLLEVVPDVPPEYYSRFRVKNLPRF 241
 Qy 241 LGSNDQDTFFTSTKRHQILFEILAKTPYGHEKKNLLGIHQLLAEGVLSAAFPLHDGPFT 300

```

Db      242  |||||LGSDNQDTFFTSTKRHQILFEILAKTPYGHEKKNLLGIHQLLAEGVLSAAFPLHDGPFKT 301
Qy      301  PPEGQPAPRLNQRQVLFQHWARWGKWNKYQPLDHVRRYFGEKVALYFAWLGFYTGWLLPA 360
Db      302  |||||PPEGQPAPRLNQRQVLFQHWARWGKWNKYQPLDHVRRYFGEKVALYFAWLGFYTGWLLPA 361
Qy      361  AVVGTLVFLVGCFLVFSDIPTQELCGSKDSFEMCPLCLDCPFWLLSSACALAQAGRLFDH 420
Db      362  |||||AVVGTLVFLVGCFLVFSDIPTQELCGSKDSFEMCPLCLDCPFWLLSSACALAQAGRLFDH 421
Qy      421  GGTVFFSLFMALWAVLLLEYWKRKSATLAYRWDCSDYEDTEERPRPQFAASAPMTAPNPI 480
Db      422  |||||GGTVFFSLFMALWAVLLLEYWKRKSATLAYRWDCSDYEDTEERPRPQFAASAPMTAPNPI 481
Qy      481  TGEDEPYFPERSRARRMLAGSVVIVMVAVVMCLVSIILYRAIMAIVVSRSGNTLLAAW 540
Db      482  |||||TGEDEPYFPERSRARRMLAGSVVIVMVAVVMCLVSIILYRAIMAIVVSRSGNTLLAAW 541
Qy      541  ASRIASLTGSVNVLFILILSKIYVSLAHVLTRWEMHRTQTKFEDAFTLKVFIFQFVNFI 600
Db      542  |||||ASRIASLTGSVNVLFILILSKIYVSLAHVLTRWEMHRTQTKFEDAFTLKVFIFQFVNFI 601
Qy      601  SSPVYIAFFKGRFVGYPGNYHTLFGVRNEECAAGGCLIELAQELLVIMVGKQVINNMQEV 660
Db      602  |||||SSPVYIAFFKGRFVGYPGNYHTLFGVRNEECAAGGCLIELAQELLVIMVGKQVINNMQEV 661
Qy      661  LIPKLGWQKQFRLRSKRRKAGASAGASQGPWEDDYELVPCEGLFDEYLEMVLQFGFVTI 720
Db      662  |||||LIPKLGWQKQFRLRSKRRKAGASAGASQGPWEDDYELVPCEGLFDEYLEMVLQFGFVTI 721
Qy      721  FVAACPLAPL FALLNNWVEIRLDARKEVCEYRRPVAERAQDIGIWFHILAGLTHLAVISN 780
Db      722  |||||FVAACPLAPL FALLNNWVEIRLDARKEVCEYRRPVAERAQDIGIWFHILAGLTHLAVISN 781
Qy      781  AFLAFSSDFLPRAYYRWTRAHDLRGFLNFTLARAPSSFAAAHNRTC 827
Db      782  |||||AFLAFSSDFLPRAYYRWTRAHDLRGFLNFTLARAPSSFAAAHNRTC 828

```

RESULT 4

ABG15488

ID ABG15488 standard; protein; 898 AA.

XX

AC ABG15488;

XX

DT 18-FEB-2002 (first entry)

XX

DE Novel human diagnostic protein #15479.

XX

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder.

XX

OS Homo sapiens.

XX

PN WO200175067-A2.

XX

PD 11-OCT-2001.

XX

PF 30-MAR-2001; 2001WO-US008631.

XX

PR 31-MAR-2000; 2000US-00540217.

PR

23-AUG-2000; 2000US-00649167.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Drmanac RT, Liu C, Tang YT;

XX

DR WPI; 2001-639362/73.

DR

N-PSDB; AAS79675.

XX

PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.

XX

[illegible]

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic amino acid sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 898 AA;

Query Match 75.5%; Score 3736; DB 4; Length 898;
Best Local Similarity 82.3%; Pred. No. 0;
Matches 727; Conservative 4; Mismatches 16; Indels 136; Gaps 6;

Qy	1	MRMAATAWAGLQGPPPLPTLCPAVRTGLYCRDQAHAE-----	37
Db	1	MRMAATAWAGLQGPPPLPTLCPAVRTGLYCRDQAHAEATDVLLAPFCQPKTRSHGTCPP	60
Qy	38	-----W---AMTSETS-----SG	47
Db	61	TERDPRGEGSTEYPRGVDGIQGWGTRALTGWTDRLLLCQACQTLPPRHWFLLPGARGWLGG	120
Qy	48	SHCA-----RSRMLRRRAQEEDSTVLIDVSPPEAEKRGSYGSTAH	87
Db	121	SPCAHQGESLSPSQSPILLRVESVKSRLMLRRRAQEEDSTVLIDVSPPEAEKRGSYGSTAH	180
Qy	88	ASEPGGQQAACRAGSPAKPRIADFVLVWEEDLKLDQQDSAARDRTDMHRTWRETFLDN	147
Db	181	ASEPGGQQAACRAGSPAKPRIADFVLVWEEDLKLDQQDSAARDRTDMHRTWRETFLDN	240
Qy	148	LRAAGLCVDQDDVDQGNTTVHYALLSASWAVLCYYAEDLRLKLPQLQELPNQASNWSAGLL	207
Db	241	LRAAGLCVDQDDVDQGNTTVHYALLSASWAVLCYYAEDLRLKLPQDYPTRPPTGRPACC	300
Qy	208	AWLGIPNVLLEVVPDVPPEYYSCRFRVNKLPRFLGSDNQDTFFTSTKRHQILFEILAKTP	267
Db	301	AWLGIPNVLLEVVPDVPPEYYSCRFRVNKLPRFLGSDNQDTFFTSTKRHQILFEILAKTP	360
Qy	268	YGHEKKNLLGIHQLLAEGVLSAAFLPHDGPFKTPPEGQAPRLNQRQVLFQHWARWGKWN	327
Db	361	YGHEKKNLLGIHQLLAEGVLSAAFLPHDGPFKTPPEGQAPRLNQRQVLFQHWARWGKWN	420
Qy	328	KYQPLDHVRRYFGEKVALLYFAWLGFYTGWLLPAAVVGTLVFLVGCFLVFSDIPTQELCGS	387
Db	421	KYQPLDHVRRYFGEKVALLYFAWLGFYTGWLLPAAVVGTLVFLVGCFLVFSDIPTQELCGS	480
Qy	388	KDSFEMCPLCLDCPFWLLSSACALAQ---AGRLFDHGGTVFFSLFMALWAVLLLEYWKR	443
Db	481	KDSFEMCPLCLDCPFWLLSSACALAQVREEAGRLFDHGGTVFFSLFMALWAVLLLEYWKR	540
Qy	444	KSATLAYRWDCSDYEDTEERPRPQFAASAPMTAPNPITGEDEPYFPERSRARRMLAGSVV	503
Db	541	KSATLAYRWDCSDYEDTEERPRPQFAASAPMTAPNPITGEDEPYFPERSRARRMLAGSVV	600
Qy	504	IVVMVAVVMCLVSIILYRAIMAVVSRSGNTLLAAWASRIASLTGSVNVLFILILSKI	563
Db	601	IVVMVAVVMCLVSIILYRAIMAVVSRSGNTLLAAWASRIASLTGSVNVLFILILSKI	660
Qy	564	YVSLAHVLTRWEMHRTQTKFEDAFTLKVFIFQVNFYSSPVYIAFFKGRFVGYPGNYHTL	623


```

Db      661 YVSLAHVLTWRWEMHRTQTKFEDAFTLKVFIFQFVNFYSSPVYIAFFKGRFVGYPGNYHTL 720
Qy      624 FGVRNEECAAGGCLIELAQELLVIMVGKQVINNMQEVLIPLKKGWWQKFRRLRSKKRKAGA 683
      |||||
Db      721 FGVRNEECAAGGCLIELAQELLVIMVGKQVINNMQEVLIPLKKGWWQKFRRLRSKKRKAGA 780
Qy      684 SAGASQGPWEDDYELVPCEGLFDEYLEM----- 711
      |||||
Db      781 SAGASQGPWEDDYELVPCEGLFDEYLEMAGFCPNACPELVPPELPEKARDQPEARSAG 840
Qy      712 -----VLQFGFVTIFVAACPLAPLFALLNNWVEIRLDARKF 747
      |||||
Db      841 QDSRPEAVLQFGFVTIFVAACPLAPLFALLNNWVEIRLDARKF 883

```

RESULT 5

ADB64420

ID ADB64420 standard; protein; 920 AA.

XX

AC ADB64420;

XX

DT 04-DEC-2003 (first entry)

XX

DE Human protein encoded by clone FEBRA20031280.

XX

KW Human; pharmaceutical; diagnostic; gene therapy; tissue regeneration;
 KW cell regeneration; membrane protein; signal transduction-related protein;
 KW transcription-related protein; osteoporosis; neurological disease;
 KW cancer; tumour.

XX

OS Homo sapiens.

XX

PN EP1308459-A2.

XX

PD 07-MAY-2003.

XX

PF 28-MAR-2002; 2002EP-00007401.

XX

PR 05-NOV-2001; 2001JP-00379298.

PR

25-JAN-2002; 2002US-00350978.

XX

PA (HELI-) HELIX RES INST.

PA (REAS-) RES ASSOC BIOTECHNOLOGY.

XX

PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
 PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
 PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;

XX

DR WPI; 2003-450961/43.

DR

N-PSDB; ADB62450.

XX

PT New polynucleotides and polypeptides, useful for developing a diagnostic
 PT marker or medicines for regulation of their expression and activity, or
 PT as targets of gene therapy.

XX

PS Claim 1; Page; 222pp; English.

XX

CC The invention discloses a polynucleotide comprising a sequence selected
 CC from 1970 fully defined nucleotide sequences which encode novel
 CC polypeptides. Also claimed is a polypeptide encoded by the polynucleotide
 CC or its partial peptide, an antibody binding to the polypeptide or peptide
 CC of the polynucleotide, immunologically assaying the polypeptide or
 CC peptide of the polynucleotide by contacting the polypeptide or peptide
 CC with the antibody of the encoded protein, and observing the binding
 CC between the two, a transformant carrying the polynucleotide in an
 CC expressible manner and an antisense polynucleotide. The oligonucleotide
 CC is useful as a primer for synthesising the polynucleotide, or as a probe
 CC for detecting the polynucleotide. The polynucleotides and encoded
 CC proteins are useful as pharmaceutical agents and many disease-related
 CC genes may be included in them, for developing a diagnostic marker or
 CC medicines for regulation of their expression and activity, or as targets
 CC of gene therapy. The genes are involved in tissue and/or cell
 CC regeneration. Membrane proteins, signal transduction-related proteins,
 CC transcription-related proteins, disease-related proteins and genes
 CC encoding them can be used as indicators for diseases (e.g. osteoporosis,
 CC neurological diseases, cancer, tumours. The cDNA may be used to regulate

CC the activity or expression of the encoded protein to treat diseases. The
CC sequence presented is a protein of the invention. Note: Some of the
CC sequence data for this patent is not represented in the printed
CC specification, but is based on sequence information supplied by the
CC European Patent Office.

XX

SQ Sequence 920 AA;

Query Match 30.9%; Score 1531.5; DB 7; Length 920;
Best Local Similarity 37.9%; Pred. No. 2.5e-148;
Matches 360; Conservative 168; Mismatches 316; Indels 105; Gaps 29;

```
Qy 44 TSSGSHCARSMRLRRRAQEEDSTVLID----VSPPEAE-----KRGSYGST---AHASEP 91
   :||| :||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 4 SSSGITNGKTKVFPVA--KDVNILFDELEAVSSPCKDDDSLHPGNLTSTSDASRLEA 61

Qy 92 GGQQAACRAGS-----PAKPRIADFLVWEEEDLKLDROQDSAARDRTDMHRTWRETFLD 146
   ||: : : :| ||| :||| : : : :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 62 GGETVPERNKSNGLYFRDGKCRI-DYILVYRK-----SNPQTEK----REVFER 105

Qy 147 NLRAAGLCVDQDQDQDGNNTTVHYALLSASWAVLCYAE DLRLKLPLQE----LPNQASNW 202
   |:| | | :||| :| :| :| | | | | | | :| :| :| :| :| :| :| :| :| :| :| :|
Db 106 NIRAEGLOMEKESSLI-NSDII FVKLHAPWEVLGRYAEQMNVRMPFRRKIYYLPRRYKFM 164

Qy 203 S-----AGLLAWLGIPNVLL--EVVPDVPP-EYYSRFRVNKLPRFLGSDNQDTFFTST 253
   | :| :| | | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 165 SRIDKQISRLRRWLPPKKPMRLDKETLPDLEENDCYTAPFSQQRHHFI-IHNKETFFNNA 223

Qy 254 KRHQILFEILAKTPYGHEKKNLGIHQLLAEGVLSAAFPLHDGPFKTPPEGPQAPRLNQR 313
   | :| :| | | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 224 TRSRIVHHILQRIKY-EEGKNKIGLNRLLTNGSYEAAFPLHEGSYRSKNSIRTHGAENHR 282

Qy 314 QVLFQHWARWGKWNKYQPLDHVRRYFGEKVALYFAWLGFYTGWLLPAAVVGTLVFLVGC 373
   :||| :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 283 HLLYECWASWGWYKYQPLDLVRRYFGEKIGLYFAWLGWYTGMLFPAAFIGLFVFLYGV 342

Qy 374 LVFSDIPTQELCGSKDSFEMCPLC-LDCPFWLLSSACALAQAGRLFDHGGTVFFSLFMAL 432
   : : :||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 343 TLDHSQVSKVEVCQATDII-MCPVCDKYCPFMRLSDSCVYAKVTHLFDNGATVFFAVFMAV 401

Qy 433 WAVLLLEYWKRKSATLAYRWDCSDYEDTEERPRPQFAAS-APMTAPNPITGEDEPYFPER 491
   || :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 402 WATVLEFVKRRRAVIAYDWDLIDWEEEEIEIRPQFEAKYSKKERMNPISGKPEPYQFT 461

Qy 492 SRARRMLAGSVVIVVMVAVVVMCLVSIILYRAIMAVVSRSGNTLLA-AWA-----SRIA 545
   : :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 462 DKCSRLIVSASGIFFMICVIAAVFGIVYRVTV-----STFAAFKVALIRNNSQVA 514

Qy 546 SLTGSVV--NLVFILILSKIYVSLAHLVTRWEMHRTQTKFEDAFTLKVFIFQFVNFYSSP 603
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 515 T-TGTAVCINFCIIMLLNVLYEKVALLLTNLEQPRTESEWENSFTLKMFLFQFVNLSST 573

Qy 604 VYIAFFKGRFVGYPGNYHTLFG-VRNEECAAGGCLIELAQELLVIMVGKQVINNMQEVLI 662
   ||||| ||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 574 FYIAFFLGRFTGHPGAYLRLINRWRL ECHPSGCLIDLQMGMIMVLKQTWNNFMELGY 633

Qy 663 PKLKGWQKFLRLSKRRKAGASAGASQGPWEDDYELVPCE--GLFDEYLEMVLQFGFVTI 720
   | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 634 PLIQNWWR---RKVRQEHGPERKISFPQWEKDYNLQPMNAYGLFDEYLEMILQFGFTTI 690

Qy 721 FVAACPLAPL FALLNNWVEIRLDARKEVCEYRRPVAERAQDIGIWFHILAGLTHLAVISN 780
   ||||| ||||| ||||| :||| ||| :||| :| ||| :||| :| ||| :||| :| ||| :||| :|
Db 691 FVAAPPLAPLLALLNNIIEIRLDAYKFVTQWRRPLASRAKDIGIWFHILAGLTHLAVISN 750

Qy 781 AFLFASFDFLPRAYRW-----TRAHDLRGFLNFTLA-----RAP 816
   ||||| :||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 751 AFVIAITSDFIPLVYAYKYGPCAGQGEAGQKCMVGYNASLSVFRISDFENRSEPEDSG 810

Qy 817 SSFAAAHNRTCRYRAFRDDDGH-----YSQTYWNLLAIRLAFVIVFEHVVSFVGRLLDLL 871
   | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 811 SEFSGTPLKYCRYRDYRDPHSLVPYGYTLQFVHVAARLAFIIVFEHLVFCIKHLISYL 870

Qy 872 VPDIPESVEIKVKREYYLAKQALAENEVLFGTNGTKDEQPKGSELSSHW 920
   :||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 871 IPDLPKDLRDRMRREKYLIQEMMYEALERLQKERKERKKNKKAHHNEW 919
```

http://es/ScoreAccessWeb/GetItem.action?AppId=10552515&seqId=775625&ItemName... 11/17/2006

Qy 140 WRETFLDNLRAAGLCVDQDDVDGNTTVHYALLSASWAVLCYYAEDLRKLPLQELPNQ- 198
 Db 159 -REEFEHNLM EAGLEL-EKDLENKSQGSIFVRIHAPWQVLAREAEFLKIKVPTKKEMYEI 216

Qy 199 -----ASNWSAGLLAWLGIPNVLLVVPDVPPEYSCRFRVKNLP-----RFLGSDNQ 246
 Db 217 KAGGSIACKFS AAL-----QKLSHLQPRV-PEHSNNKMKNLSYPFSREKMYLYNIQEK 269

Qy 247 DTFFTSTKRHQILFEILAKTPYGHEKKNLLGIHQLLAEGVLSAAFPLHDGPFKTPPEGPQ 306
 Db 270 DTFFDNATRSRIVHEILKRTACS-RANNTMGINSLIANNIYEAAYPLHDGEYDSPEDD-- 326

Qy 307 APRLNQRQVLFQHWARWGKWNKYQPLDHRVRYFGEKVALYFAWLGFYTGWLLPAAVVGT 366
 Db 327 ---MNDRLKLYQEWARYGVFYKFQPIDLRKYFGEKIGLYFAWLGLYTSFLIPSSVIGVI 383

Qy 367 VFLVGCFLVFSDIPTQELCGSKDSFEMCPLC-LDCPFWLLSSACALAQAGRLFDHGGTVF 425
 Db 384 VFLYGCATIEEDI PSREMCDDQNAFTMCPLCDKSCDYWNLSACGTAQASHLFDNPATVF 443

Qy 426 FSLFMALWAVLLEWKRKSATLAYRWDCSDYEDTEER----PRPQFAA----- 470
 Db 444 FSIFMALWATMFLENWKRLQMRGLGYFDLTGIEEEERAEQHSRPEYETKVREKMLKESN 503

Qy 471 -SAPMTAPNPIT----GEDEPYFPERSRARRMLAGSVVIVVMVAVVMCLVSIILYRAIM 525
 Db 504 QSAVQKLENTNTECGDEDEDKLTWKDRFPGYLMNFASILFMIALTFSIVFGVIVYRITT 563

Qy 526 AIVVSRSGNTLLAAWASRI---ASLTGSVVNLVFIILSKIYVSLAHVLRWEMHRTQTK 582
 Db 564 AAALS-----LNKATRSNVRVTVTATAVIINLVVILILDEIYGAVAKWLTKIEVPKTEQT 618

Qy 583 FEDAFTLKVFIFQVNFYSSPVYIAFFKGRFVGPNGYHTLF-GVRNEECAAGGCLIELA 641
 Db 619 FEERLILKAFLLKFVNAYSPIFYVAFKGRFVGRPGSYVYVFDGYRMEECAPGGCLMELC 678

Qy 642 QELLVIMVGKQVI-NNMQEVLIPKLKGWQKFRLRSKKRKAGASAGA-SQGP--WEDDYE 697
 Db 679 IQLSIIMLGKQLIQNNIFEIGVPKLK---KLFRKLKDETEAGETSAHSKHPEQWDLDS 735

Qy 698 LVPCEGLFDEYLEMVLQFGFVTIFVAACPLAPL FALLNNWVEIRLDARKFVCEYRRPVAE 757
 Db 736 LEPYTGLTPEYMEMIIQFGFVTLFVASFP LAPVFALLNNVIEVRLDAKKFVTELRRPDAV 795

Qy 758 RAQDIGIWFHILAGLTHLAVISNAFLAFSSDFLPRAYRWTRAHD--LRGFLNFTLA-- 813
 Db 796 RTKDIGIWFIDILSGIGKFSVISNAFVIAITSDFIPRLVYQYSYSHNGTLHG FVNHTLSFF 855

Qy 814 -----RAPSSFAAAHNRTCryAFRD-----DDGHYSQTYWLLAIRLAFVIVFEH 859
 Db 856 NVSQLKEGTQ PENSQFDQEVQFCRFKDYREPPWAPNPYEFKQYWFILSARLAFVIFQN 915

Qy 860 VVFSVGRLLDLLVPDIPESVEIKVRE-----YYLAKQALAENEVLFGTNGTKDEQPKG 913
 Db 916 LVMFLSVLVDWMIPDIPDISQIKKEKSLVDFFLKE---EHEKCLKMDEPALRSPGG 971

Qy 914 SELSSHWTPFTVPAK-SQL 931
 Db 972 GDRSRSRRAASSAPSGSQL 990

RESULT 11

ABB62812

ID ABB62812 standard; protein; 1219 AA.

XX

AC ABB62812;

XX

DT 26-MAR-2002 (first entry)

XX

DE Drosophila melanogaster polypeptide SEQ ID NO 15228.

XX

KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.

XX

OS Drosophila melanogaster.

XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US009231.
 XX
 PR 23-MAR-2000; 2000US-0191637P.
 PR 11-JUL-2000; 2000US-00614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR WPI; 2001-656860/75.
 DR N-PSDB; ABL06915.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signaling and cell-cell
 PT interactions.
 XX
 PS Disclosure; SEQ ID NO 15228; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
 CC ABB72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 1219 AA;

Query Match 29.2%; Score 1445; DB 4; Length 1219;
 Best Local Similarity 35.6%; Pred. No. 3.5e-139;
 Matches 342; Conservative 165; Mismatches 332; Indels 122; Gaps 27;

Qy	35	AERWAMTSETSSGSHCARSRML-----RRRAQEEDSTVLIDVSPPEAEKRGSY	82
		: : : : :	
Db	249	ADRVNQSYEVMESSH---SNVLPDQFGYRQLIPTERKASDTASSV-----SGSY	294
Qy	83	GSTAHASEP---GGQAAACRAGSPAKP-----RIADFVLVW-EEDLKLDRQ	125
		: : : : : : :	
Db	295	YGSRKASKSNSLGGESGDERRVSKQDREGLDPESLMFRDGRKRVDMVLAWEEEDLGMTE	354
Qy	126	QDSAARDRTDMHRTWRETFDNLRAAGLCVDQDQD-VQDGNTTVHYALLSASWAVLCYYAE	184
		: : : : : : : : :	
Db	355	AEAKRRDN-----RRSFMENLIKEGLEVELEDKSQSFNEKTFFLKIHLPWRLETRLAE	407
Qy	185	DLRLKLP-----LQELPNQASNWSAGLLAWLGIPNVLLEVVPDVP	225
		: : : : : : :	
Db	408	VMNLKLPVKRFITISVKPSWDEENVVLRNMQYWKDVWQR-LTKKIQLDQTLLE----GET	462
Qy	226	EYYSRFRVKNLPRFLGSDNQDFTFTSTKRHQILFEILAKTPYGHEKKNLLGIHQLLAEG	285
		: : : : : : : : : : :	
Db	463	TFKATANGNPEEQFIVKD-RATAFTSAQRSLMVMQVLIRTPFDESDRS--GIRRLMNDG	519
Qy	286	VLSAAFPLHDGPFKTPPEGPQAPRLN-QRQVLFQHWARWGKWNKYQPLDHVRRYFGEKVA	344
		: : : : : : : :	
Db	520	TYLGCFPLHEGRY----DRPHSSGISLDRRVLYQTWAHP SQWYKKQPLCLVRKYFGDKIA	575
Qy	345	LYFAWLGFYTGWLLPAAVVGTLVFLVGCFLVFS--IPTQELCG--SKDSFEMCPLC-LD	399
		: : : : :	
Db	576	LYFCWLGFYTEMLVYPVAVGTLCFIYGLATLESDNTPSKEICNEYGTGNITLCPLCDKA	635
Qy	400	CPFWLLSSACALAQAGRLFDHGGTVFFSLFMALWAVLLLEYWKRKSATLAYRWDCSDYED	459
		: : : : : : : :	
Db	636	CSYQRLSESCFLSRLTYLFDNPSTVFFAIFMSFWATTFLELWKRKQSVLVWEWDLHNV-D	694
Qy	460	TEERPRPQFAASAPMTAPNPITGEDEPYFPERSRARRMLAGSVVIVVMVAVVVMCLVSII	519
		: : : : : : : : :	
Db	695	MDEENRPEFETNATTFRMNPVTREKEPYMSTWNR SIRFVITGSAVLFMISVVL SAVLGTI	754

Qy 520 LYRAIMAIIVVSRSGNTLLAAWASRIASLTGSSVNLVFIILSKIYVSLAHVLTRWEMHRT 579
 ||| : | : | : | :| ::||| |::||| :| || | ||
 Db 755 LYRITLVSVIYGGGGFFVKEHAKLFTSVTAALINLVVIMILTRIYHRMAIKLTNLENPRT 814
 Qy 580 QTKFEDAFTLKVFIFQFVNFIYSSPVYIAFFKGRFVGYPGNYHT----LFGVRNEECAAGG 635
 |::||::| |::| |::| |||| :||| |||| |||: |::| | |
 Db 815 HTEYEDSYTFKIFFFEFNMFYSSLIYIAFFKGRFFDYPGDDQARKSEFFRLKNDICDPAG 874
 Qy 636 CLIELAQELLVIMVGKQVINNMQEVLPKLGWWQKFRLSKKRKAGASAGASQGPWEDD 695
 || || :| :||| || || || || ||: :| :|| || |
 Db 875 CLSELCIQLAIIIMVGKQCWNNFMEYLFPKFWNWWR----QRKHKQATKDESHLMMAWEQD 930
 Qy 696 YELV-PCE-GLFDEYLEMVLQFGFVTIFVAACPLAPLALLNNWVEIRLDARKFVCEYRR 753
 | : | |||||::|::|::| ||||| ||||| | | : ||
 Db 931 YHMQDPGRALFDEYLEMILQYGFVTIFVAAFPLAPLALLNNVAEIRLDAYKMTQARR 990
 Qy 754 PVAERAQDIGIWFHILAGLTHLAVISNAFLAFSSDFLPRAYR--WTRAHDLRGFLNFT 811
 |::|| :||| | : || :| :||::|::|::|::| | : :| | : :
 Db 991 PLAERVEDIGAWYGILRIITYTAVVSNAFVIAYTSDFIPRMVYKFVYSEHTLAGYIEHS 1050
 Qy 812 LA-----RAPSSFAAANRTCRYRAFRDDDGHY-----SPTYWNLALAIRLAFVIVF 857
 | : :| : ||::| :| | |::| | ||::| | ||||::||
 Db 1051 LSIFNTSDYKEEWGASVSEKDPDTCQYRGYRNGPKDYEYGLSPHYWHVFAARLAFVVF 1110
 Qy 858 EHVVSFVGRLLDLLVPDIPESVEIKVKREYYLAKQALAENEVLFGTNGTKDEQPKGSELS 917
 ||||| : : :||:| | : ::|| ||::| : :| | | : :
 Db 1111 EHVVFVITGIMQFIIPDVPSEVKTQMQRQLLAKEAKYQ-----HGKRAQGDSQDIM 1163
 Qy 918 S 918
 |
 Db 1164 S 1164

RESULT 12

ADC42854

ID ADC42854 standard; protein; 910 AA.

XX

AC ADC42854;

XX

DT 18-DEC-2003 (first entry)

XX

DE REMAP protein #14.

XX

KW Cytostatic; Antiarteriosclerotic; Anti-HIV; Antiinflammatory;

KW Antiallergic; Antidiabetic; REMAP; pathogenesis.

XX

OS Homo sapiens.

XX

PN WO2003027228-A2.

XX

PD 03-APR-2003.

XX

PF 16-JUL-2002; 2002WO-US022833.

XX

PR 17-JUL-2001; 2001US-0306020P.

PR 27-JUL-2001; 2001US-0308179P.

PR 02-AUG-2001; 2001US-0309702P.

PR 10-AUG-2001; 2001US-0311476P.

PR 10-AUG-2001; 2001US-0311551P.

PR 10-AUG-2001; 2001US-0311718P.

PR 24-AUG-2001; 2001US-0314798P.

PR 31-AUG-2001; 2001US-0316639P.

PR 07-SEP-2001; 2001US-0317996P.

XX

PA (INCY-) INCYTE GENOMICS INC.

XX

PI Lal PG, Honchell CD, Forsythe IJ, Walia NK, Tang TY, Borowsky ML;

PI Barroso I, Yue H, Warren BA, Thangavelu K, Gietzen KJ, Azimzai Y;

PI Lee EA, Baughn MR, Gorvad AE, Duggan BM, Tran B, Li JX;

PI Richardson TW, Elliott VS, Zebbarjadian Y, Tran UK, Yao MG;

PI Peterson DP, Luo W, Lehr-Mason PM;

XX

DR WPI; 2003-421156/39.

XX

PT New human receptors and membrane-associated proteins (REMAP), useful for
 PT diagnosing, treating or preventing disorders associated with aberrant

PT REMAP expression, e.g. cancer, AIDS, atherosclerosis, hypertension or
PT stroke.
XX
PS Claim 1; SEQ ID NO 14; 115pp; English.
XX
CC The present invention relates to an isolated polypeptide. The
CC polypeptides and polynucleotides are useful in diagnosing, treating and
CC preventing disorders associated with aberrant expression of REMAP, such
CC as cell proliferative, autoimmune/inflammatory, renal, neurological,
CC cardiovascular, metabolic, developmental, endocrine, muscle,
CC gastrointestinal, lipid metabolism or transport disorders, and viral
CC infections. These are also useful in assessing the effects of exogenous
CC compounds on the expression of nucleic acids and amino acid sequences of
CC REMAP, in facilitating drug discovery process, and in investigating the
CC pathogenesis of diseases or medical conditions. Expression and
CC purification were achieved using bacterial or virus-based expression
CC systems. The present sequence represents an REMAP protein of the
CC invention.
XX
SQ Sequence 910 AA;

Query Match 28.3%; Score 1402.5; DB 7; Length 910;
Best Local Similarity 38.2%; Pred. No. 5.7e-135;
Matches 322; Conservative 157; Mismatches 286; Indels 79; Gaps 24;

```

Qy 108 RIADFVLVWEEDLKLDQRQDSAADRTDMHRTWRETFLDNLRAAGLCVD-QQDVQDGNNT 166
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 67 RRIDFVLVYED----ESRKETNKKGTNEKQRRKQAYESNLICHGLQLEATRSVLDDKL 122

Qy 167 VHYALLSASWAVLCYYAEDLRKLPLQELPNQASNSA--GLLAWLGIPNVLLVVPDVP 224
   : : : : | | | | | | | | | | | | | | | | | | | | | | | |
Db 123 --FVKVHAPWEVLCTYAEIMHIKPLK--PNDLKNRSSAFGTLNWFVKVLSVDESIIKPE 178

Qy 225 PEYYSRFRVKNLPRFLGSDNQDTFFTSTKRHQILFEILAKTPYG-HEKKNLLGIHQLLA 283
   | : : | | : | | : | | : | : | : | : | : | : | : | : | : |
Db 179 QEFTTAPFEKNRMNDIFYVD-RDAFFNPATRSRIVYFILSRVKYQVINNVSKFGINRLVN 237

Qy 284 EGVLSAAFLHDGPFKTPPEGPQAPRLNQRQVLFQHWARWGKWNKYQPLDHVRRYFGEKV 343
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 238 SGIYKAAFPLHDCKFRQSEDPSCP--NERYLLYREWAHPRSIYKKQPLDLIRKYYGEKI 295

Qy 344 ALYFAWLGFYTGWLLPAVVVGTLVFLVGCFLVFSDIPTQELC----GSKDSFEMCPLCLD 399
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 296 GIYFAWLGYTQMLLLAAVVGACFLYGYLNQDNCTWSKEVCHPDIGGK--IIMCPQC-D 352

Qy 400 --CPFWLLSSACALAQAGRLFDHGGTVFFSLFMALWAVLLLEYWKRKSATLAYRWDCSDY 457
   | | | | | : | : : : | | | | | : | : | : | : | | | | | :
Db 353 RLCPFWKNLITCESSKKLCIFDSFGTLVFAVFMGVWVTLFLEFWKRRQAELEYEWDTVEL 412

Qy 458 EDTEERPRPQFAASAPMTAPNPITGEDE--PYFPERSRARRMLAGSVVIVVMVAVVMCL 515
   : | | : | : | | | | | | | | | | | | | | | | | | | | : : :
Db 413 QQ-EEQARPEYEARTHVINEITQEEERIPFTAWGKCIRITLCASAVF-FWILLIIASV 470

Qy 516 VSIILYRAIMAIIVSR-----SGNTLLAAWAS--RIASLTGSVNVLFILILSKIYVSL 567
   : | | : | : | | | | | | | | | | | | | | | | | | | | :
Db 471 IGIIVYRLSVFIVFSAKLPKNINGTDPIQKYLTPTATTSITASIISFIIIMILNTIYEKV 530

Qy 568 AHVLTWRWEMHRTQTKFEDAFTLKVFIFQFVNFYSSPVYIAFFKGRFVGYPGN-YHTLFGV 626
   | : : | : : | | | | : : : | : : : | | | | | | | | | | | :
Db 531 AIMITNFELPRTQTDYENSLTMKMFLLQFVNYSSCFYIAFFKGFVGPDPVYWLKGY 590

Qy 627 RNEECAAGGCLIELAQELLVIMVGKQVINNMQEVLI PKLKGWWQKFRRLRSKKRKAGASAG 686
   | | | | | | | | | | : | : | | : | : | | | | : | : | | |
Db 591 RNEECDPGGCLLELTQTIIIMGGKAIWNNIQEVLLPWIMNLIGRFHRVSGSEKITPR-- 648

Qy 687 ASQGPWEDDYELVPCE--GLFDEYLEMVLQFGFVTIFVAACPLAPL FALLNNWVEIRLDA 744
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 649 -----WEQDYHLQPMGKLGFLFYEYLEMIIQFGFVTLFVASFPLAPLLALVNNILEIRVDA 703

Qy 745 RKFVCEYRRPVAERAQDIGIWFHILAGLTHLAVISNAFLAFSSDFLPRAYRW----- 799
   | : : | | | : | | | | | | : | : | | : | : | | : | : | |
Db 704 WKLTQFRLVPEKAQDIGAWQPIMQGIAILAVVTNAMIIFTSDMIPRLVYYSFVSVP 763

Qy 800 ----RAHDLRGFLNFTLARAPSSFAAA-----HNRTCRYRAFRDDDDGH-- 838
   : : : | : | | | | | | | | | | | | | | | | | | | | |
Db 764 YGDHTSYTMEGYINNTL----SIFKVADFKNKSKGNPYSDLGNHHTTCRYRDFRYPPGHPQ 819

```

Qy 839 ---YSQTYWNLLAIRLAFVIVFEHVSVGRLLDLLVDPDIPESVEIKVKREYYLAKQALA 895
 Db 820 EYKHNIYYWHVIAAKLAFIIVMEHVIYSVKFFISYAIPDVSKRTKSKIQREKYLTQKLLH 879

Qy 896 ENEV 899
 Db 880 ENHL 883

RESULT 13

ABB65993

ID ABB65993 standard; protein; 1075 AA.

XX

AC ABB65993;

XX

DT 26-MAR-2002 (first entry)

XX

DE Drosophila melanogaster polypeptide SEQ ID NO 24771.

XX

KW Drosophila; developmental biology; cell signalling; insecticide;

KW pharmaceutical.

XX

OS Drosophila melanogaster.

XX

PN WO200171042-A2.

XX

PD 27-SEP-2001.

XX

PF 23-MAR-2001; 2001WO-US009231.

XX

PR 23-MAR-2000; 2000US-0191637P.

PR 11-JUL-2000; 2000US-00614150.

XX

PA (PEKE) PE CORP NY.

XX

PI Venter JC, Adams M, Li PWD, Myers EW;

XX

DR WPI; 2001-656860/75.

DR N-PSDB; ABL10096.

XX

PT New isolated nucleic acid detection reagent for detecting 1000 or more

PT genes from Drosophila and for elucidating cell signaling and cell-cell

PT interactions.

XX

PS Disclosure; SEQ ID NO 24771; 21pp + Sequence Listing; English.

XX

CC The invention relates to an isolated nucleic acid detection reagent

CC capable of detecting 1000 or more genes from Drosophila. The invention is

CC useful in developmental biology and in elucidating cell signalling and

CC cell-cell interactions in higher eukaryotes for the development of

CC insecticides, therapeutics and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA

CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-

CC ABB72072). The sequence data for this patent did not form part of the

CC printed specification, but was obtained in electronic format directly

CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX

SQ Sequence 1075 AA;

Query Match 27.7%; Score 1369.5; DB 4; Length 1075;
 Best Local Similarity 37.4%; Pred. No. 1.9e-131;
 Matches 313; Conservative 163; Mismatches 283; Indels 77; Gaps 20;

Qy 108 RIADFVLVWEEDLKLDQRQDSAAARDRTDMHRT-WRETFLDNLRAAGLCVD--QQDVQDGN 164
 Db 192 RSIDFVLAYRIN-----AHEPTELENTEKRRVFANLISQGLEVESSQKD----- 236

Qy 165 TTVHYALLSASWAVLCYYAEDRLRLKPLQLPNQASNWSAGLLAWLGIPNVL-----LE 218
 Db 237 -QIWFKIHAPLEVLRRYAEILKLRMPMKEIPGMSVVNRSTKSVFSSLKHVFQFFLRNIY 295

Qy 219 VVPDPVPEYYSCRFRV--NKLPRFLGSDNQDTFFTSTKRHQILFEIL--AKTPYGHEKKN 274
 Db 296 VDEEIFPK-RAHRTAIYSRDKEYLFDIRQDCFFTTAVRSRIVEFILDQRFPKQNHDM 354

```

Qy      275 LLGIHQLLAEGVLSAAFLHDGPFKTPPEGPAAPRLNQRQVLFQHWARWGKWNKYQPLDH 334
Db      355 AFGIERLIAEGVYSAAAYPLHDGEITETG-----TMRALLYKHWASVPKWYRYQPLDD 406

Qy      335 VRRYFGKEKVALYFAWLGFYTGWLLPAAVVGTFLVFLVGCFLVFSDIPTQELCGS-KDSFEM 393
Db      407 IKEYFGVKIGLYFAWLGYTYMLLLASIVGVICFLYSWFSKKNYVPVKDICQSGNTNITM 466

Qy      394 CPLCLDCPFWLLSSACALAQAGRLFDHGGTVFFSLFMALWAVLLLEYWKRKSATLAYRWD 453
Db      467 CPLCDWCNFWDLKETCNYAKVTYLIDNPSTVFFAVFMSFWATLFLELWKRYSAEITHRWD 526

Qy      454 CSDYEDTEERPRPQFAA---SAPMTAPNPITGEDEPYFP-ERSRARRMLAGSVVIVVMVA 509
Db      527 LTGFVHVEEHRPQYLARLEHIPPTRVDYVTNIKEPTVPFWRMKLPATVFSFSVLLLIA 586

Qy      510 VVVMCLVSIIILYRAIMAVVSRSGNTLLAAWASRIASLTGSVNVLFILILSKIYVSLAH 569
Db      587 LAFVALLAVVYRMSMLAALKVGASPMITSSAIVLATASAAFNLCCLLYILNMYNHLAE 646

Qy      570 VLTRWEMHRTQTKEFADFTLKVFIFQFVNFYSSPVYIAFFKGRFVGYPGNYHTLFGVRNE 629
Db      647 YLTELEMWRTQTQFDDSLTLKIYLLQFVNYIASIFYIAFFKGRFVGHPGEYNKLFDRQE 706

Qy      630 ECAAGGCLIELAQELLVIMVGKQVINNMQEVLPKLGWWQKFRLRSKKRKAGASAGASQ 689
Db      707 ECSSGGCLTELICQLAIMVGKQAFNTILEVYLP---FWRKV---LAIQVGLSRLFNN 759

Qy      690 GP-----WEDDYELVP--CEGLFDEYLEMVLQFGFVTIFVAACPLAPLFFALLNNV 738
Db      760 TPNPDKAKDERWMRDFKLLDWGARGLFPEYLEMVLQYGFVTIFVAAFPLAPFFALLNNIL 819

Qy      739 EIRLDARKFVCEYRRPVAERAQDIGIWFHILAGLTHLAVISNAFLAFSSDFLPRAYRW 798
Db      820 EMRLDAKLLTHHKRPVSQRVRDIGVWYRILDCIGKLSVITNGFIIAFTSDMIPRLVYRH 879

Qy      799 --TRAHDLRGFLNFTLAR-----APSSFAAAHN---RTCRYRAFR-----DDGHSYQT 842
Db      880 YVNKQGTLDGYLNFTLSEFKVIDSPTLYSLAGDLSNITTCRYTDFRLPPSSPEKYTLSSM 939

Qy      843 YWNLLAIRLAFVIVFEHVVSFVGRLLDLLVPDIPESVEIKVKREYLLAKQALAENE 898
Db      940 FYIILACRLGFVVVFENFVALVMILVRWCIPDMSVELRDQIRREVYVTNEIIDIQ 995

```

RESULT 14

AEG11145

ID AEG11145 standard; protein; 712 AA.

XX

AC AEG11145;

XX

DT 20-APR-2006 (first entry)

XX

DE Human transmembrane protein 16A, SEQ ID NO: 10.

XX

KW Genetic marker; diagnostic; prognosis; gastrointestinal tumor;
cytostatic; neoplasm; tumor marker; transmembrane protein 16A.

XX

OS Homo sapiens.

XX

PN US2006040292-A1.

XX

PD 23-FEB-2006.

XX

PF 08-JUL-2005; 2005US-00177894.

XX

PR 08-JUL-2004; 2004US-0586676P.

XX

PA (WEST/) WEST R B.

PA (VRIJ/) VAN DE RIJN M.

XX

PI West RB, Van De Rijn M;

XX

DR WPI; 2006-182760/19.

DR N-PSDB; AEG11140.

DR GENBANK; AAH27590.

XX
PT Classifying tumor as gastrointestinal stromal tumor belonging to PDGFRA
PT positive subclass, involves detecting expression or activity of gene
PT encoding DOG1 polypeptide in sample.
XX
PS Disclosure; SEQ ID NO 10; 177pp; English.
XX
CC The present invention relates to three gene markers such as DOG1, KIT and
CC platelet derived-growth factor receptor alpha (PDGFRA) that are useful in
CC classifying tumors. These gene markers are useful in the classification
CC of gastrointestinal stromal tumors (GISTs) and tumors other than GISTs.
CC The invention also relates to methods providing diagnostic, prognostic
CC and predicative information based on the classifying step. The invention
CC is useful for classifying gastrointestinal stromal tumors as belonging to
CC a PDGFRA positive subclass, KIT negative or PDGFRA negative subclass. The
CC present sequence is human transmembrane protein 16A (DOG1; TMEM16A). The
CC DOG1 gene encodes a transmembrane protein of unknown function
CC (transmembrane protein 16A). The transmembrane protein 16A is encoded by
CC DOG1 gene that is mapped to 11q13.2 on chromosome 11.
XX
SQ Sequence 712 AA;

Query Match 27.6%; Score 1367.5; DB 10; Length 712;
Best Local Similarity 41.6%; Pred. No. 1.6e-131;
Matches 299; Conservative 128; Mismatches 220; Indels 71; Gaps 17;

```

Qy      259 LFEILAKTPYGHEKKNLLGIHQLLAEGVLSAAFPLHDGPFKTPPEGPQAPRLNQRQVLFO 318
      :||| :| : :|| ||| || :||:|||| : : :| :||:
Db      2 VYEILKRTTCTKAKYS-MGITSLLANGVYAAAYPLHDGDY-----NGENVEFNDKLLYE 55

Qy      319 HWARWGKWNKYQPLDHVRRYFGEKVALYFAWLGFYTGWLLPAAVVGTLLVFLVGCFLVFS 378
      |||:| : |||:| ||:||||: ||||| || |:|:|:| :||| : :
Db      56 EWARYGVFYKYQPIDLVRKYFGEKIGLYFAWLGVYTQMLIPASIVGIIIFLYGCATMDEN 115

Qy      379 IPTQELCGSKDSFEMCPLC-LDCPFWLLSSACALAQAGRLFDHGGTVFFSLFMALWAVLL 437
      ||: |:| : : |||| | :| :|||| |:| |||: ||||:|||||
Db      116 IPSMEMCDQRHNTMCPLCDKTCYWKMSACATARASHLFDNPATVFFSVFMALWAATF 175

Qy      438 LEYWKRSATLAYRWDCSDYEDTEE----RPRPQFAA-----SAPMTAPNPITGEDEPYF 488
      |:||| | ||| : :| || || :| | : | | :
Db      176 MEHWKRKQMLNRYRWDLTGFEETEEAVKDHPRAEYEARVLEKSLKKESRNKET--DKVKL 233

Qy      489 PERSRARRMLAGSVVIVVMVAVVMCLVSIILYRAIMAVVSRSGNTLLAAWASRIASLT 548
      | | | | :| :|| : :||| || : : : : : :
Db      234 TWRDRFPAYLTNLVSIIFMIAVTFIVLGVIIYRISMAAALAMNSSPSVRSNIRVTVTAT 293

Qy      549 GSVVNLVFIILSKIYVSLAHVLTWRWEMHRTQTKEFADFLLKVFIFQFVNFYSSPVYIAF 608
      :||| ||:| :| :| ||: |: :| :| :| :||| |: |:|
Db      294 AVIINLVVILLDEVYGCIARWLTKEIVPKTEKSFEERLIFKAFLLKFVNSYTPIFYVAF 353

Qy      609 FKGRFVGYPGNYHTLF-GVRNEECAAGGCLIELAQELLVIMVGQVI-NMQEVLIPKLK 666
      ||||| ||:| :| | ||| |||:| :| :||:|:| ||: |:| :|
Db      354 FKGRFVGRPGDYVYIFRSFRMEECAPGGCLMELCIQLSIIMLGKQLIQNNLFEIGIPKMK 413

Qy      667 GWWQKFRRLRSKKRKAGASAGASQGPWEDDYELVPCGLFDEYLEMVLQFGFVTIFVAACP 726
      : :| : : : :| || | || ||:|:|:| |||:|
Db      414 KLIRYLKLKQQSPPDHEECVKKRQRYEVDYNLEPFAGLTPEYMEMIIQFGFVTLFVASFP 473

Qy      727 LAPLFA LLNNWVEIRLDARKFVCEYRRPVAERAQDIGIWFHILAGLTHLAVISNAFLAF 786
      ||||| ||| :|||:| ||| ||:| |||:| || : ||| |||:|
Db      474 LAPLFA LLNNIIEIRLDAKKFVTELRRPVAVRAKDGIWYNILRGIGKLAVIINAFVISF 533

Qy      787 SSDFLPRA--YYRWTRAHDLRGFLNFTLARAPSSF-----AAAHN-----RTCR 828
      :||:| | :| : :|:| || || ||:|:| ||| : ||
Db      534 TSDFIPRLVLYLYMSKNGTMHGFVNHTL----SSFNVSDFGNGTAPNDPLDLGYEVQICR 589

Qy      829 YRAFRD---DDGHY--SQTYWNLLAIRLAFVIVFEHVSVGRLLDLLVPDIPESVEIKV 883
      |: :| : | |: :| :| |||||:| : :| :||| : :
Db      590 YKDYREPPWSENKYDISKDFWAVLAARLAFVIVFQNLVFMMSDFVDWVDPIDPKDISQEI 649

Qy      884 KREYYLA-----KQALAEVLFGTNGTKDEQP-----KGSELSSH 919
      :| | | ||| | || |
Db      650 HKEKVLMLVELFMREEQDKQQLL--ETWMEKERQKDEPPCNHNTKACPDLSLGSPAPSH 705

```

RESULT 15


```

ABB65022
ID   ABB65022 standard; protein; 1058 AA.
XX
AC   ABB65022;
XX
DT   26-MAR-2002 (first entry)
XX
DE   Drosophila melanogaster polypeptide SEQ ID NO 21858.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US009231.
XX
PR   23-MAR-2000; 2000US-0191637P.
PR   11-JUL-2000; 2000US-00614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC, Adams M, Li PWD, Myers EW;
XX
DR   WPI; 2001-656860/75.
DR   N-PSDB; ABL09125.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signaling and cell-cell
PT   interactions..
XX
PS   Disclosure; SEQ ID NO 21858; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC   ABB72072). The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1058 AA;

```

Qy	108	RIADFVLVWEEDLKLDRQQDSAARDRTDMHRTWRETFLDNLRAAGLCVDQDQDVQDGNTTV	167
		: : : : : : :	
Db	204	RSVDFVLAYNGETQLEE-----HRRKCEIFEANLQREGLQLEHNKVQ----RV	247
Qy	168	HYALLSASWAVLCYYAEDLRLLKLPLQELPNQASNWSAGLLAWLGIPNVILLEVPDPPEY	227
		: : : : : : : : : :	
Db	248	HFIKIHPAEVLYRYAEILKIKVPLKPIPGQD-----QIFAESAEHF	289
Qy	228	YSCRFRVNKL-----PRF-----LGSNDQDTFFTSTKRHQILFE	261
		: : : : : : : : :	
Db	290	KTCFSRMCKSLFSSVQLNTALFPEREPRIHLEFSRNYLELYDTEHPNFLDASTRYSIINF	349
Qy	262	ILAKTPY--GHEKKNLGIHQLLAEGVLSAAFLHDGPFKTPPEGPQAPRLNQ-----RQ	314
		: : : : : : : : : : :	
Db	350	ILQRQRFVEGEETADNLGIEKLVQDGVYTCAYTLHDVERRSRSAAKGVGQHIQVEEQLRE	409
Qy	315	VLFQHWARWGKWNKYQPLDHVRRYFGEKVALYFAWLGFYTGWLLPAAVVGTLVFLVGCFL	374
		: : : : :	
Db	410	TLKPFYC-----SLQPLDQIKDYFGAKVALYFAWLGFYTQMLIPISVFGVLCFLYGFIT	463
Qy	375	VFSDIPTQELCGSKDSFEMCPLC-LDCPFWLLSSACALAAAGRLFDHGGTVFFSLFMALW	433
		: : : : : : : : : : :	
Db	464	WNSDPISRDICDDNGTI-MCPOCDRSCDYWRNLNETCTSSKFNYLIDNNMTVVFAFSAIWI	522

Qy 434 AVLLLEYWKRKSATLAYRWDCSDYEDTEERPRPQFAASAPMTAPNPITGEDEPYFPERSR 493
 Db 523 AVVYLEFWKRYASAGLVHRWGLTGFTHHVEHPRPQYLARISRT--KKLAG--KAYEQDHTG 578

Qy 494 ARRLAGSV-----VIVVMVAVVMCLVSIILYRAIMAVVSRSGNTLLA 538
 Db 579 KRTILDPPVFPWSFKFLPNFTSYSIMVLFICISVIAIAGIIYR-----MAQRASHSILG 633

Qy 539 AWAS-----RIASLTGSVVNLVFIILSKIYVSLAHVLRWEMHRTQTKFEDAFTLKVFI 593
 Db 634 SENSMTFKMILPMTAGIIDLIVISLDMVYSNLAVKLTNYEYCRTQTEYDESLTIKNYV 693

Qy 594 FQFVNFYSSPVYIAFFKGRFVGYPGNHYHTLFGVRNEECAAGGCLIELAQELLVIMVGKQV 653
 Db 694 FQFVNYSSLFYIAFLKGKFGVYPAKYNRVLGFRQEENPGGCLMELCMQLVIIMAGKQA 753

Qy 654 INNMQEVLIPKL-----KGWW---QKFRRLRSKKRKAGASAGASQGPWEDDYELVP 700
 Db 754 VNAIVEMLIPYLMRTFKELSYRHGWYKSHQDQRL-----VPYNQFTEDYNLLP 801

Qy 701 CE--GLFDEYLEMVLQFGFVTIFVAACPLAPLFAALLNNWVEIRLDARKEVCEYRRPVAER 758
 Db 802 AENNSLYVEYLEMVMVQFGFITLFSLAFPLALLNNVIEVRLDAIKMLRFLRRPVGMR 861

Qy 759 AQDIGIWFHILAGLTHLAVISNAFLAFSSDFLPRAYYRWTRAH-DLRGFLNFTLA---- 813
 Db 862 ARDIGVWHSIMTVVTRIAVASSAMIIAFSTNLIPKIVYAASMGDPENNYLNFTLAVFNT 921

Qy 814 ----RAPSSFAAAH-NRT-CRYRAFRD--DDGH-YSQ--TYWNLLAIRLAFVIVFEHVVF 862
 Db 922 KDFVQVQLLGGSQHVNETVCRYTEFRNSPEDPHYPKRPMIYWKILTGRLAFIVIQNIIT 981

Qy 863 SVGRLLDLLVPDIPESVEIKVKREYYLAKQALAENE 898
 Db 982 MLQGILRWAVPDVSGRLLKRIKRENFLLREHIIEYE 1017

Search completed: October 27, 2006, 20:23:21
 Job time : 214 secs

SCORE 1.3 BuildDate: 12/06/2005

SCORE Search Results Details for Application 10552515 and Search Result us-10-552-515-1.ra1.

[Score Home](#) [Retrieve Application](#) [SCORE System](#) [SCORE](#) [Comments /](#)
[Page](#) [List](#) [Overview](#) [FAQ](#) [Suggestions](#)

This page gives you Search Results detail for the Application 10552515 and Search Result us-10-552-515-1.ra1.

[start](#)

[Go Back to previous page](#)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2006, 20:29:09 ; Search time 54 Seconds
(without alignments)
1512.335 Million cell updates/sec

Title: US-10-552-515-1
Perfect score: 4950
Sequence: 1 MRMAATAWAGLQGPPLPTLC.....SELSSHWTPFTVPKASQLQQ 933

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_AA:*
1: /EMC_Celerra_SIDS3/ptodata/2/iaa/5_COMB.pep:*
2: /EMC_Celerra_SIDS3/ptodata/2/iaa/6_COMB.pep:*
3: /EMC_Celerra_SIDS3/ptodata/2/iaa/7_COMB.pep:*
4: /EMC_Celerra_SIDS3/ptodata/2/iaa/H_COMB.pep:*
5: /EMC_Celerra_SIDS3/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /EMC_Celerra_SIDS3/ptodata/2/iaa/RE_COMB.pep:*
7: /EMC_Celerra_SIDS3/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1531.5	30.9	920	2	US-10-104-047-2574	Sequence 2574, Ap
2	1154	23.3	596	2	US-10-104-047-2541	Sequence 2541, Ap
3	912.5	18.4	475	2	US-10-104-047-3116	Sequence 3116, Ap
4	796	16.1	425	2	US-09-270-767-45552	Sequence 45552, A
5	396.5	8.0	215	2	US-09-270-767-61064	Sequence 61064, A
6	353	7.1	366	2	US-09-270-767-32253	Sequence 32253, A
7	353	7.1	366	2	US-09-270-767-47470	Sequence 47470, A
8	290	5.9	189	2	US-09-270-767-31816	Sequence 31816, A
9	290	5.9	189	2	US-09-270-767-47033	Sequence 47033, A
10	255.5	5.2	199	2	US-09-270-767-31722	Sequence 31722, A

11	255.5	5.2	199	2	US-09-270-767-46939	Sequence 46939, A
12	186.5	3.8	166	2	US-09-621-976-4064	Sequence 4064, Ap
13	117	2.4	548	1	US-08-676-279-50	Sequence 50, Appl
14	117	2.4	548	2	US-08-903-139B-8	Sequence 8, Appli
15	117	2.4	548	2	US-08-637-823B-25	Sequence 25, Appl
16	117	2.4	548	2	US-09-614-957D-25	Sequence 25, Appl
17	115.5	2.3	2013	1	US-08-324-977-12	Sequence 12, Appl
18	115.5	2.3	2013	1	US-08-384-616-12	Sequence 12, Appl
19	115.5	2.3	2013	1	US-08-904-686A-12	Sequence 12, Appl
20	115.5	2.3	2013	2	US-09-315-850-12	Sequence 12, Appl
21	115.5	2.3	3010	1	US-08-324-977-2	Sequence 2, Appli
22	115.5	2.3	3010	1	US-08-324-977-14	Sequence 14, Appl
23	115.5	2.3	3010	1	US-08-384-616-2	Sequence 2, Appli
24	115.5	2.3	3010	1	US-08-384-616-14	Sequence 14, Appl
25	115.5	2.3	3010	1	US-08-904-686A-2	Sequence 2, Appli
26	115.5	2.3	3010	1	US-08-904-686A-14	Sequence 14, Appl
27	115.5	2.3	3010	2	US-09-315-850-2	Sequence 2, Appli
28	115.5	2.3	3010	2	US-09-315-850-14	Sequence 14, Appl
29	110.5	2.2	680	2	US-09-725-735A-19	Sequence 19, Appl
30	108.5	2.2	523	2	US-09-949-016-11540	Sequence 11540, A
31	108.5	2.2	578	2	US-09-052-753B-7	Sequence 7, Appli
32	106	2.1	539	2	US-09-248-796A-16542	Sequence 16542, A
33	105	2.1	1089	2	US-10-012-231A-102	Sequence 102, App
34	105	2.1	1089	2	US-10-015-389A-102	Sequence 102, App
35	105	2.1	1089	2	US-10-006-768A-102	Sequence 102, App
36	105	2.1	1089	2	US-10-015-671A-102	Sequence 102, App
37	105	2.1	1089	2	US-10-015-393A-102	Sequence 102, App
38	105	2.1	1089	2	US-10-011-833A-102	Sequence 102, App
39	105	2.1	1089	2	US-10-006-041A-102	Sequence 102, App
40	105	2.1	1089	2	US-10-012-064A-102	Sequence 102, App
41	105	2.1	1089	2	US-10-015-392A-102	Sequence 102, App
42	105	2.1	1089	3	US-10-011-795B-102	Sequence 102, App
43	105	2.1	1089	3	US-10-015-386A-102	Sequence 102, App
44	105	2.1	1089	3	US-10-012-121A-102	Sequence 102, App
45	105	2.1	1089	3	US-10-006-485A-102	Sequence 102, App

ALIGNMENTS

RESULT 1

US-10-104-047-2574
 ; Sequence 2574, Application US/10104047
 ; Patent No. 6943241
 ; GENERAL INFORMATION:
 ; APPLICANT: HELIX RESEARCH INSTITUTE
 ; TITLE OF INVENTION: No. 6943241el full length cDNA
 ; FILE REFERENCE: H1-A0105
 ; CURRENT APPLICATION NUMBER: US/10/104,047
 ; CURRENT FILING DATE: 2002-03-25
 ; PRIOR APPLICATION NUMBER:
 ; PRIOR FILING DATE:
 ; NUMBER OF SEQ ID NOS: 4096
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2574
 ; LENGTH: 920
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-104-047-2574

Query Match 30.9%; Score 1531.5; DB 2; Length 920;
 Best Local Similarity 37.9%; Pred. No. 2e-157;
 Matches 360; Conservative 168; Mismatches 316; Indels 105; Gaps 29;

Qy	44	TSSGSHCARSMRLRRRAQEEDSTVLID----VSPPEAE-----KRGSYGST---AHASEP	91
		: : : : : : : : : :	
Db	4	SSSGITNGKTKVFHPVA--KDVNILFDELEAVSSPCKDDSLHPGNLTSTSDASRLEA	61
Qy	92	GGQQAACRAGS-----PAKPRIADFLVWEEDLKLDQRQDSAARDRTDMHRTWRETFLD	146
		: : : :: : : : :	
Db	62	GGETVPERNKSNGLYFRDGKCRI-DYILVYRK-----SNPQTEK----REVFER	105
Qy	147	NLRAAGLCVDQDVQDGNNTVHYALLSASWAVLCYAE DLRLKLPLQE----LPNQASNW	202
		: : : : : : : : : :	
Db	106	NIRAEG LQMEKESSLI-NSDIIFVKLHAPWEVLGRYAEQMNVRMPFRRKIYYLP RRYKFM	164

Qy 357 LLPAAVVGTIVFLVGCFLVFSDIPTQELCGSKDSFEMCPLC-LDCPFWLLSSACALAQAG 415
 | | | | : | | | | : : | | : | | | | | : | :
 Db 2 LFPAAFIGLFLVFLYGVTTLDHSQVSKEVCQATDII-MCPVCDKYCPFMRLSDSCVYAKVT 60

Qy 416 RLFDHGGTVFFSLFMALWAVLLLEYWKRKSATLAYRWDCSDYEDTEERPRPQFAAS-APM 474
 | | | : | | | : | | | : | | | : | | | : | | | : | | | :
 Db 61 HLFNDGATVFFAVFMAVWATVFLFVWKRRAVIAYDWDLDWEEEEEEIRPQFEAKYSKK 120

Qy 475 TAPNPITGEDEPYFPERSRARRMLAGSVVIVVMVAVVMCLVSIILYRAIMAIIVVSRSGN 534
 | | | : | | | : | | : | | : | | : | | : | | : | | :
 Db 121 ERMNPISGKPEPYQAFDTKCSRILVSASGIFFMICVVIAAVFGIVYRVVTV-----S 173

Qy 535 TLLA-AWA-----SRIASLTGSVV--NLVFILILSKIYVSLAHVLRWEMHRTQTKFEDA 586
 | | | | : | | : | | : | | : | | : | | : | | : | | : | | :
 Db 174 TFAAFKWLIRNNSQVAT-TGTAVCINFCIIMLLNVLYEKVALLLTNLEQPRTESEWENS 232

Qy 587 FTLKVFIFQFVNFYSSPVYIAFFKGRFVGYPGNYHTLFG-VRNEECAAGGCLIELAQELL 645
 | | | : | | | | | | | | | | | | | | | | | | | | | | : :
 Db 233 FTLKMFLEFQFVNLNSSTFYIAFFLGRFTGHPGAYLRLINRWRLEECHPSGCLIDLQMOMG 292

Qy 646 VIMVGKQVINNMQEVLPKLGWQKFRRLRSKKRKAGASAGASQGPWEDDYELVPCE--G 703
 : | | | | | | : | : | : | : | : | : | : | : | : | : | : | :
 Db 293 IIMVLKQVWNNFMELGYPLIQNWTR---RKVRQEHGPERKISFPQWEKDYNLQPMNAYG 349

Qy 704 LFDEYLEMVLQFGFVTIFVAACPLAPLFAALLNNWVEIRLDARKEVCEYRRPVAERAQDIG 763
 | | | | | : | | | | | | | | | | | | | | | | | : | | : | | : | | :
 Db 350 LFDEYLEMILQFGFTTIFVAAFPLAPLLALLNNIEIRLDAYKFVTQWRRPLASRAKDIG 409

Qy 764 IWFHILAGLTHLAVISNAFLAFSSDFLPRAYRW-----TRAHDLRGFLNFTLA 813
 | | : | | : | : | | | : | : | | | : | : | | | : | : | | : | :
 Db 410 IWYGILEGIGILSVITNAFVIAITSDFIPRLVYAYKYGPCAGQGEAGQKCMVGYVNASLS 469

Qy 814 -----RAPSSFAAAHNRTCARYAFRDDGHH-----YSQTYWNLLAIRLAFV 854
 | | : | : | : | | : | | : | | : | | : | | : | | : | | : | | :
 Db 470 VFRISDFENRSEPESDGSEFSGTPLKYCRYRDYRDPHSLVPYGYTLQFVHVLAAARLAFI 529

Qy 855 IVFEHVVSFVGRLLDLLVPDIPESVEIKVKREYLLAKQALAENEVLFGTNGTKDEQPKGS 914
 | | | | : | : | | : | : | | : | : | | : | : | | : | : | | :
 Db 530 IVFEHLVFCIKHLISYILPDLPKDLRDRMRREKYLIQEMMYEALERLQKERKERKKNKG 589

Qy 915 ELSSHW 920
 : |
 Db 590 AHHNEW 595

RESULT 3

US-10-104-047-3116
 ; Sequence 3116, Application US/10104047
 ; Patent No. 6943241
 ; GENERAL INFORMATION:
 ; APPLICANT: HELIX RESEARCH INSTITUTE
 ; TITLE OF INVENTION: No. 6943241e1 full length cDNA
 ; FILE REFERENCE: H1-A0105
 ; CURRENT APPLICATION NUMBER: US/10/104,047
 ; CURRENT FILING DATE: 2002-03-25
 ; PRIOR APPLICATION NUMBER:
 ; PRIOR FILING DATE:
 ; NUMBER OF SEQ ID NOS: 4096
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 3116
 ; LENGTH: 475
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-104-047-3116

Query Match 18.4%; Score 912.5; DB 2; Length 475;
 Best Local Similarity 38.0%; Pred. No. 3.4e-90;
 Matches 202; Conservative 89; Mismatches 143; Indels 97; Gaps 12;

Qy 430 MALWAVLLLEYWKRKSATLAYRWDCSDYEDTEERPRPQFAASAPMTAPNPITGEDEPYFP 489
 | | : | : | | : | | | | : | : | : | : | : | : | : | : | : | :
 Db 1 MAVWATVFLFVWKRRAVIAYDWDLDWEEEE----- 32

Qy 490 ERSRARRMLAGSVVIVVMVAVVMCLVSIILYRAIMAIIVVSRSGNTLLA-AWA-----SR 543
 : | : : | : | : : | | | : | : | : | : | : | : | : | : | : | :
 : | : : | : | : : | | | : | : | : | : | : | : | : | : | : | :


```

Qy      219 VVPDVPPEYYSCRFRV--NKLPRFLGSDNQDTFFTSTKRHQILFEIL--AKTPYGHEKKN 274
      |  :: | :  : ||      : :  | | ||| :  | :| : ||  : |  ::
Db      197 VDEEIFPK-RAHRFTAIYSRDKEYLFDIRQDCFFTTAVRSRIVEFILDQRFPKQNDHDM 255

Qy      275 LLGIHQLLAEGVLSAAFPPLHDGPFKTPPEGPQAPRLNQRQVLFQHWARWGKWNKYQPLDH 334
      || :| :||| ||| :||| | :| :||| || :||| |
Db      256 AFGIERLIAEGVYSAAYPLHDGEITETG-----TMRALLYKHVASVPKWYRYQPLDD 307

Qy      335 VRRYFGKEVALYFAWLGFYTGWLLPAAVVGTLVFLVGCFLVFSDIPTQELCGS 387
      :: ||| | : ||||| :|| || :||| : ||  | : : :| : :| |
Db      308 IKEYFGVKIGLYFAWLGYTTYMLLLASIVGVICFLYSWFSLKNYVPVKDICQS 360

```

RESULT 7

```

US-09-270-767-47470
; Sequence 47470, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 47470
; LENGTH: 366
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-47470

```

```

Query Match          7.1%; Score 353; DB 2; Length 366;
Best Local Similarity 32.4%; Pred. No. 3.6e-29;
Matches 95; Conservative 56; Mismatches 104; Indels 38; Gaps 9;

```

```

Qy      108 RIADFVLVWEEDLKLDRQQDSAARDRTDMHRT-WRETFLDNLRAAGLCVD--QQDVQDGN 164
      | ||| :  :      | : || : | | | || : || | : | :|
Db      93 RSIDFVLAYRIN-----AHEPTELENTEKRRVFANLISQGLEVESSQKD----- 137

Qy      165 TTVHYALLSASWAVLCYYAEDLRLKLPLQELPNQASNWSAGLLAWLGIPNVL-----LE 218
      : : : | || ||| | :|| :|| :| : : : :| : : : :|
Db      138 -QIWFVKIHAPLEVLRRAEILKLRMPMKEIPGMSVVNRSTKSVSLSKHVFQFFLRNIY 196

Qy      219 VVPDVPPEYYSCRFRV--NKLPRFLGSDNQDTFFTSTKRHQILFEIL--AKTPYGHEKKN 274
      |  :: | :  : ||      : :  | | ||| :  | :| : ||  : |  ::
Db      197 VDEEIFPK-RAHRFTAIYSRDKEYLFDIRQDCFFTTAVRSRIVEFILDQRFPKQNDHDM 255

Qy      275 LLGIHQLLAEGVLSAAFPPLHDGPFKTPPEGPQAPRLNQRQVLFQHWARWGKWNKYQPLDH 334
      || :| :||| ||| :||| | :| :||| || :||| |
Db      256 AFGIERLIAEGVYSAAYPLHDGEITETG-----TMRALLYKHVASVPKWYRYQPLDD 307

Qy      335 VRRYFGKEVALYFAWLGFYTGWLLPAAVVGTLVFLVGCFLVFSDIPTQELCGS 387
      :: ||| | : ||||| :|| || :||| : ||  | : : :| : :| |
Db      308 IKEYFGVKIGLYFAWLGYTTYMLLLASIVGVICFLYSWFSLKNYVPVKDICQS 360

```

RESULT 8

```

US-09-270-767-31816
; Sequence 31816, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 31816
; LENGTH: 189
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-31816

```

```

Query Match          5.9%; Score 290; DB 2; Length 189;

```

Qy	249	FFTSTKRHQILFEILAKTPY---GHEKKNLLGIHQLLAEGVLSAAFPLHDGPFKTPPEGPQ	306
		:: : : : : : : : :	
Db	39	FLDASTRYSIIINFILQRQRFVEGEETADNLGIEKLVQDGVYTCAYTLHD-----	87
Qy	307	APRLNQRVQLFQHWARWGKWNKYQPLDHVRRYFGEKVLYFAWLGFYTGWLLPAAVVGLT	366
		: : : : :	
Db	88	---KDDRDRLLKEWANISKWKNLQPLDQIKDYFGAKVLYFAWLGFYTMQLPIPVFVGL	144
Qy	367	VFLVGCFLVFSDIPTQELCGSKDSFEMCPLC-LDCPFWLLSSAC	409
		: :: : : : : :	
Db	145	CFLYGFIWNSDPISRDICDDNGTI-MCPOCDRSCDYWRLNETC	187

```

US-09-270-767-47033
; Sequence 47033, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 47033
; LENGTH: 189
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-47033

```

Qy	249	FFTSTKRHQILFEILAKTPY--GHEKKNLLGIHQLLAEGVLSAAFPLHDGPFKTPPEGPQ	306
		:: : : : : : : : :	
Db	39	FLDASTRYSIIINFILQRQRFVEGEETADNLGIEKLVDQGVYTCAYTLHD-----	87
Qy	307	APRLNQRQVLFGHWRARWGKWNKYQLDHWRRYFGEKVALYFAWLGFYTGWLLPAAVVGT	366
		: : : : :	
Db	88	---KDDRDRLLKEWANISKWKNLQPLDQIKDYFGAKVALYFAWLGFYTMPLIPISVFGVL	144
Qy	367	VFLVGCFLVFSDIPTQELCGSKDSFEMCPLC-LDCPFWLLSSAC	409
		: : : : : :	
Db	145	CFLYGFIITWNSDPISRDICDDNGTI-MCPOCDRSCDYWRLNETC	187

```

US-09-270-767-31722
; Sequence 31722, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 31722
; LENGTH: 199
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-31722

```

Oy: 417 LFDHGGTVFFSLFMALWAVLLLEYWKRSATLAYRWDCSDYEDTEERPRPQFAASAPMTA 476
 | :| |: || |:|||: ||:|| | | : || : : | |||: | |
Dd: 11 LI DNNTVVAFESMAI W VVYLEFWKRY SAGLVHRWG LTGTETH HVEHP RPOYLARI SR T- 69

```

Qy      477 PNPITGEDEPYFPERSRARRMLAGSV-----VIVVMVAVVVMCLVSIILY 521
      : | : | : : | : | : : : : : : : : : | : |
Db      70 -KKLAG--KAYEQDQTGKRTILDPDPVFWFSKFLPNFTSYSIMVLFICISVIAIAGIIY 126

Qy      522 RAIMAIVVSRSGNTLLAAWAS-----RIASLTGSVVNLVFILILSKIYVSLAHVLRWEM 576
      | : | : : : | : | : | : : : | : | : | : |
Db      127 R-----MAQRASHSILGSENSMTFKVMILPMTAGIIDLIVISLLDMVYSNLAVKLTNYEY 181

Qy      577 HRTQTKFEDAFTLKVFIF 594
      ||||: : : : | : | : : |
Db      182 CRTQTEYDESLTIKNYVF 199

```

RESULT 11

```

US-09-270-767-46939
; Sequence 46939, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 46939
; LENGTH: 199
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-46939

```

```

Query Match          5.2%; Score 255.5; DB 2; Length 199;
Best Local Similarity 29.8%; Pred. No. 5.7e-19;
Matches 59; Conservative 46; Mismatches 64; Indels 29; Gaps 5;

```

```

Qy      417 LFDHGGTVFFSLFMALWAVLLLEYWKRKSATLAYRWDCSDYEDTEERPRPQFAASAPMTA 476
      | | : | | : | : | : | : | : | : | : | : | : |
Db      11 LIDNMTVVFAFSMAIWAVVYLEFWKRYAGLVHRWGLTGFTHHVEHPRPQYLARISRT- 69

Qy      477 PNPITGEDEPYFPERSRARRMLAGSV-----VIVVMVAVVVMCLVSIILY 521
      : | : | : : | : | : : : : : : : : : | : |
Db      70 -KKLAG--KAYEQDQTGKRTILDPDPVFWFSKFLPNFTSYSIMVLFICISVIAIAGIIY 126

Qy      522 RAIMAIVVSRSGNTLLAAWAS-----RIASLTGSVVNLVFILILSKIYVSLAHVLRWEM 576
      | : | : : : | : | : | : : : | : | : | : |
Db      127 R-----MAQRASHSILGSENSMTFKVMILPMTAGIIDLIVISLLDMVYSNLAVKLTNYEY 181

Qy      577 HRTQTKFEDAFTLKVFIF 594
      ||||: : : : | : | : : |
Db      182 CRTQTEYDESLTIKNYVF 199

```

RESULT 12

```

US-09-621-976-4064
; Sequence 4064, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 4064
; LENGTH: 166
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -58..-1
US-09-621-976-4064

```

Query Match 3.8%; Score 186.5; DB 2; Length 166;
Best Local Similarity 28.0%; Pred. No. 1.5e-11;
Matches 45; Conservative 34; Mismatches 75; Indels 7; Gaps 3;

```
Qy      448 LAYRWDCSDYEDTEERPRPQFAASAPMTAPNPITGEDEPYFPERSRARRMLAGSVVIVM 507
      : ||| : | ||| : | |||::|| :| | : | : | :
Db      1 MTYRWGTLMKRKFEPRPGFHG---VLGINSITGKEEPLPSYKRQLRIYLVSLPFVCL 57

Qy      508 VAVVMCLVSIILYRAIMAIIVSRSGNTLLAAWASRIASLTGSVNLVLFILILSKIYVSL 567
      : :: | : | : : : | | : |:: : | |:::|
Db      58 CLYFSLYVMMIYFDMEVWALGLHENS---SEWTS-VLLYVPSIIYAIVIEIMNRLYRYA 113

Qy      568 AHVLRWEMHRTQTKFEDAFTLKVFIFQFVNFYSSPVYIAF 608
      | || || || :: :: | || :| | :| |||
Db      114 AEFTLSWENHRLESAYQNHILKVLVFNFLNCFASLFYIAF 154
```

RESULT 13

US-08-676-279-50
; Sequence 50, Application US/08676279
; Patent No. 5869247
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: MACROPHAGE NUCLEOTIDE SEQUENCE
; NUMBER OF SEQUENCES: 63
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/676,279
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB95/00095
; APPLICATION NUMBER: GB 9400929.7
; FILING DATE: 19-JAN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9422021.7
; FILING DATE: 31-OCT-1994
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 548 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-676-279-50

Query Match 2.4%; Score 117; DB 1; Length 548;
Best Local Similarity 20.2%; Pred. No. 0.0048;
Matches 136; Conservative 89; Mismatches 207; Indels 240; Gaps 34;

```
Qy      67 VLIDVSPPEAEKRGSYGSTAHASEPG-GQQAACRAGSPAKPRIADFLVWWEEDLKLDRO 125
      :: | ||| : ||| : ||| | |||
Db      1 MISDKSPRL-SRPSYSGSI--SSLPGPAPQAPCR-----ETYLSEKIP 41

Qy      126 QDSAARDRTDMHRTWRET-----FLD--NLRAAGLCVDQDQDVGNTTVHYALLSA 174
      || : : : | | ||| | : : ||| |
Db      42 IPSADQGTFSRLKLWAF TGPF LMSIAFLDPGNI-----ESDLQAGAVAGFKLLWVL 93

Qy      175 SWA----VLC-----YEAEDLRLKL-----PLQELPNQ 198
      || : || : || : || : || : || :
Db      94 LWATVLGLLCQRLAARLG VVTGKDLGEVCHLYYPKVRILLWLTIELAIVGSDMQEVIGT 153

Qy      199 ASNW---SAGLLAWLGIPNVLLVVPDVPPEYYSCRFRVNKLPRFLGSDNQDTFFTSTKR 255
      | : : ||| : | ||: || : : : || |
Db      154 AISFNLLSAGRIPLWG--GVLITIV-DTFFFLFDNYGLRKLEAFFG----- 197

Qy      256 HQILFEILAKTPYGHEKKNLLGIHQLLAEGVLSAAFPLHDGPFKTPPEGQPAPRLNQRQV 315
      : | | : | : | : | : | | | | : :
Db      198 --LLITIMALT-FGYE---YVVAHP--SQGALLKGLVLPCTPGCGQPELLQAVGIVGAI 249

Qy      316 LFQHWARWGKWNKYQPLDHVRRYFGEKVALYFAWLGFYTGWLLPAAVVGTLVFLVGCFL- 374
      : : | : : | | : : || : | : : : || :
Db      250 MPHNIYLSALVKSREVDRTRRVDVREANMYF-----LIEATIALSVSFIINLFVM 300
```

```

Qy      375 -VFSDIPTQELCGSKDSFEMCPCLDCPFWLLSSACALAQAGRLFD-----HGG 422
      ||  | :  :::| :|  :  | :  |  ||
Db      301 AVFGQAFYQQT--NEEAFNIC-----ANSSLQNYAKIFPRDNNTVSVDIYQGG 346

Qy      423 TVFFSLF----MALWAVLLLEYWKRKSATLAY-----RWDCSDYEDTEERPRP 466
      :  ||  :  :||| ||  :  :  |  |  ||
Db      347 VILGCLFGPAALYIWAVGLLAAGQSSTMTGTIYAGQFVMEGFLKLRW----- 392

Qy      467 QFAASAPMTAPNPITGEDEPYFPERSR-ARRMLAGSVVIV--VMVAV-----V 511
      || || :|  |  | :  | :|||  :
Db      393 -----SRFARVLLTRSCAILPTVLVAVFRDLKDLGLNDL 427

Qy      512 VMCLVSIILYRAIMAIVVSRSGNTLLAAWAS-RIASLTGS-----VNLVLFILILSKI 563
      :  | :| :|  | :|  | :  :| :|  |  :|| :| :
Db      428 LNLVQLSLLPFAVLPIITFTSMPAVMQEFANGRMSKAITSCIMALVCAINLYFVI---SY 484

Qy      564 YVSLAHVLTREWEMHRTQTKFEDAFTLKVFIFQFVNFYSSPVYIAF-----FKGRFVGYPG 618
      || |  | :  | :|  :  :  | :| :  | :
Db      485 LPSLPH-----PAYFGLVALFA-IGYLGLTAYLAWTCCIAHGATFLTHSS 528

Qy      619 NYHTLFGVRNEE 630
      :  | :| :| |||
Db      529 HKHFLYGLPNEE 540

```

RESULT 14

US-08-903-139B-8

; Sequence 8, Application US/08903139B

; Patent No. 6114118

; GENERAL INFORMATION:

; APPLICANT: Joe W. Templeton, Jianwei Feng, L. Garry Adams,
 ; APPLICANT: Erwin Schurr, Philippe Gros, Donald S. Davis and Roger Smith
 ; TITLE OF INVENTION: METHOD OF IDENTIFICATION OF ANIMALS
 ; TITLE OF INVENTION: RESISTANT OR SUSCEPTIBLE TO DISEASES SUCH AS RUMINANT
 ; TITLE OF INVENTION: BRUCELOSIS, TUBERCULOSIS, PARATUBERCULOSIS AND SALMONELLOSIS
 ; NUMBER OF SEQUENCES: 31

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Pravel, Hewitt, Kimball & Krieger
 ; STREET: 1177 West Loop South, 10th Floor
 ; CITY: Houston
 ; STATE: TX
 ; COUNTRY: USA
 ; ZIP: 77027-9095

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/903,139B

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 60/031,443

; FILING DATE: September 20, 1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Krieger, Paul E.
 ; REGISTRATION NUMBER: 25,886
 ; REFERENCE/DOCKET NUMBER: 00162-3/V96171US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 713-850-0909
 ; TELEFAX: 713-850-0165

; INFORMATION FOR SEQ ID NO: 8:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 548 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

US-08-903-139B-8

Query Match 2.4%; Score 117; DB 2; Length 548;

Best Local Similarity 20.2%; Pred. No. 0.0048;

Matches 136; Conservative 89; Mismatches 207; Indels 240; Gaps 34;

```

Qy      67 VLIDVSPPEAEKRGSYGSTAHASEPG-GQQAACRAGSPAKPRIADFLVWEEDLKLDQRQ 125
      :: | |||  | |||  :| ||  | |||  |

```

Db 1 MI\$DKSPRL-SRPSYGS-SSLPGPAPQAPCR-----ETYLSEKIP 41

Qy 126 QD\$AARDRTMDHRTWRET-----FLD--NLRAAGLCVDQDQDQDNTTVHYALLSA 174
 || : : : | | || | : : : | |

Db 42 IPSADQGTFSRLKLWAF TGPGFLMSIAFLDPGNI-----ESDLQAGAVAGFKLLWVL 93

Qy 175 SWA----VLC-----YYAEDLRLKL-----PLQELPNQ 198
 || : || : || : || : || : ||

Db 94 LWATVGLLCCQRLAARLGVVTKDLGEVCHLYYPKVPRILLWLTIELAIVGSDMQEVI GT 153

Qy 199 ASNW---SAGLLAWLGIPNVLLEVDPVPEYYSRFRVKNLPRFLGSDNQDTFFTSTKR 255
 | : : || : | || : | : : : || |

Db 154 AISFNLLSAGRIPLWG--GVLITIV-DTFFFLFDNYGLRKLEAFFG----- 197

Qy 256 HQILFEILAKTPYGHEKNLLGIHQLLAEGVLSAAFPLHDGPFKTPPEGPQAPRLNQRQV 315
 : | | : | : | : | : | : | : | : |

Db 198 --LLITIMALT-FGYE---YVVAHP--SQGALLKGLVLP TCPCGCGQPELLQAVGIVGAI I 249

Qy 316 LFQHWARWGKWKYQPLDHVRRYFGEKVALYFAWLGFYTGWLLPAAVVGTLVFLVGCFL- 374
 : : : : | : | : | : | : | : | : |

Db 250 MPHNIYLSALVKSREVDTRRVDVREANMYF-----LIEATIALSVSFIINLFVM 300

Qy 375 -VFSDIPTQELCGSKDSFEMCPLCLDCPFWLLSSACALAQAGRLFD-----HGG 422
 || : : : | : : : | : : |

Db 301 AVFGQAFYQQT--NEEAFNIC-----ANSSLQNYAKIFPRDNNTVSVDIYQGG 346

Qy 423 TVFFSLF----MALWAVLLLEYWKRKSATLAY-----RWDCSDYEDTEERPRP 466
 : | : : | : : | : : | : : |

Db 347 VILGCLFGPAALYIWA VGLLAAGQSSTMTGT YAGQFVMEGFLKLRW----- 392

Qy 467 QFAASAPMTAPNPITGEDEPYFPERSR-ARRMLAGSVVIV--VMVAV-----V 511
 || || : | | : : | : | : |

Db 393 -----SRFARVLLTRSCAILPTVLVAVFRDLKDL SGLNDL 427

Qy 512 VMCLVSIILYRAIMAI VVSRSNTLLAAWAS-RIASLTGS-----VNLV FILILSKI 563
 : | | : | : | : : : | : : : | : : : |

Db 428 LNVLSLLLPFAVLPILTFTSMPAVMQEFANGRMSKAITSCIMALVCAINLYFVI---SY 484

Qy 564 YVSLAHVLTREWHRHTQTKFEDAFTLKVFIQFVN FYSSPVYIAF-----FKGRFVGYPG 618
 || | : | : | : : | : : | : : |

Db 485 LPSLPH-----PAYFGLVALFA-IGYLGLTAYLAWTCCIAHGATFLTHSS 528

Qy 619 NYHTLFGVRNEE 630
 : | | : | |

Db 529 HKHFLYGLPNEE 540

RESULT 15

US-08-637-823B-25

; Sequence 25, Application US/08637823B

; Patent No. 6184031

; GENERAL INFORMATION:

; APPLICANT: Gros, Philippe

; APPLICANT: Skamene, Emil

; TITLE OF INVENTION: DNA SEQUENCES THAT ENCODE A NATURAL

; TITLE OF INVENTION: RESISTANCE TO INFECTION WITH INTRACELLULAR PARASITES

; NUMBER OF SEQUENCES: 32

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: KLAUBER & JACKSON

; STREET: 411 Hackensack Ave

; CITY: Hackensack

; STATE: New Jersey

; COUNTRY: U.S.A.

; ZIP: 07601

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/637,823B

; FILING DATE: 05/08/96

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Jackson, David A.

; REGISTRATION NUMBER: 26,742

```
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487 5800
; TELEFAX: 201 343 1684
; TELEX:
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 548 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-637-823B-25
```

```
Query Match      2.4%; Score 117; DB 2; Length 548;
Best Local Similarity 20.2%; Pred. No. 0.0048;
Matches 136; Conservative 89; Mismatches 207; Indels 240; Gaps 34;
```

```
Qy      67 VLIDVSPPEAEKRGSYGSTAHASEPG-GQAAACRAGSPAKPRIADFLVWVEEDLKLDRO 125
      :: | ||| | |||| :| || | ||| | |
Db      1 MISDKSPRL-SRPSYGSI--SSLPGPAPQPAPCR-----ETYLSEKIP 41

Qy     126 QDAAARDRTDMHRTWRET-----FLD--NLRAAGLCVDQDQDQDNTTVHYALLSA 174
      || : : : | | ||| | : : | : | |
Db     42 IPSADQGTFSRLKLWAF TGPGFLMSIAFLDPGNI-----ESDLQAGAVAGFKLLWVL 93

Qy     175 SWA----VLC-----YYAEDLRLKL-----PLQELPNQ 198
      || :|| : || : | : | : || :
Db     94 LWATVLGLLCQRLAARLGVVTGKDLGEVCHLYPKVPRILLWLTIELAIVGSDMQEVIGT 153

Qy     199 ASNW---SAGLLAWLGINVLLLEVVPDPPEYYSCRFRVKNLPRFLGSDNQDTFFTSTKR 255
      | : : ||| : | || : | | : : || | |
Db     154 AISFNLLSAGRIPLWG--GVLITIV-DTFFFLFDNYGLRKLEAFFG-----197

Qy     256 HQILFEILAKTPYGHEKKNLLGIHQLLAEGVLSAAFPLHDGPFKTPPEGPQAPRLNQRV 315
      :| | : | : | : | : | | | | | : :
Db     198 --LLITIMALT-FGYE---YVVAHP--SQGALLKGLVLP TCPCGCGQPELLQAVGIVGAI 249

Qy     316 L FQHWARWGKWNKYQPLDHVRRYFGEKVALYFAWLGFYTGWLLPAAVVGTLVFLVGCFL- 374
      : : : | : | | : || : | : | : | : | :
Db     250 MPHNIYLSALVKSREVDRTTRVDVREANMYF-----LIEATIALSVSFIINLFVM 300

Qy     375 -VFSDIPTQELCGSKDSFEMCPLCLDCPFWLLSSACALAQAGRLFD-----HGG 422
      || | : : : | : | : : | : | : |
Db     301 AVFGQAFYQQT--NEEAFNIC-----ANSSLQNYAKIFPRDNNTVSVDIYQGG 346

Qy     423 TVFFSLF---MALWAVLLLEYWKRKSATLAY-----RWDCSDYEDTEERPRP 466
      : | : : ||| || : : | | ||
Db     347 VILGCLFGPAALYIWA VGLLAAGQSSTMTGT YAGQFVMEGFLKLRW-----392

Qy     467 QFAASAPMTAPNPITGEDEPYFPERSR-ARRMLAGSVVIV--VMVAV-----V 511
      || || :| | | : | : ||| :
Db     393 -----SRFARVLLTRSCAILPTVLVAVFRDLKDL SGLNDL 427

Qy     512 VMCLVSIILYRAIMAI VVSRSGNTLLAAWAS-RIASLTGS-----VNLVFIILSKI 563
      : | | : | | : | : : | : | : | : | :
Db     428 LNVLSLLL PFAVLPI LFTSMPAVMQEFANGRMSKAITSCIMALVCAINLYFVI---SY 484

Qy     564 YVSLAHVLRWEMHRTQTKFEDAFTLKVFIFQFVN FYSSPVYIAF-----FKGRFVGYPG 618
      || | : | : | : | : | : | : | :
Db     485 LPSLPH-----PAYFGLVALFA-IGYLGLTAYLAWTCCIAHGATFLTHSS 528

Qy     619 NYHTLFGVRNEE 630
      : | | : | |
Db     529 HKHFLYGLPNEE 540
```

```
Search completed: October 27, 2006, 20:30:36
Job time : 57 secs
```

SCORE 1.3 BuildDate: 12/06/2005

SCORE Search Results Details for Application 10552515 and Search Result us-10-552-515-1.rapbm.

[Score Home](#) [Retrieve Application](#) [SCORE System](#) [SCORE](#) [Comments /](#)
[Page](#) [List](#) [Overview](#) [FAQ](#) [Suggestions](#)

This page gives you Search Results detail for the Application 10552515 and Search Result us-10-552-515-1.rapbm.

[start](#)

[Go Back to previous page](#)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2006, 20:29:57 ; Search time 189 Seconds
(without alignments)
2286.664 Million cell updates/sec

Title: US-10-552-515-1
Perfect score: 4950
Sequence: 1 MRMAATAWAGLQGPPPLPTLC.....SELSSHWPFTVPKASQLQQ 933

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2097797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters: 2097797

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA_Main:*
1: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
3: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
4: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
5: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
6: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	3736	75.5	898	5 US-10-450-763-45847	Sequence 45847, A
2	1531.5	30.9	920	4 US-10-104-047-2574	Sequence 2574, Ap
3	1531.5	30.9	920	6 US-11-072-512-2574	Sequence 2574, Ap
4	1488	30.1	960	6 US-11-177-894-7	Sequence 7, Appli
5	1479.5	29.9	840	6 US-11-177-894-11	Sequence 11, Appl
6	1445	29.2	1219	6 US-11-097-143-15228	Sequence 15228, A
7	1402.5	28.3	910	5 US-10-484-148-14	Sequence 14, Appl
8	1369.5	27.7	1075	6 US-11-097-143-24771	Sequence 24771, A
9	1367.5	27.6	712	6 US-11-177-894-10	Sequence 10, Appl
10	1199.5	24.2	1058	6 US-11-097-143-21858	Sequence 21858, A
11	1154	23.3	596	4 US-10-104-047-2541	Sequence 2541, Ap

12	1154	23.3	596	6	US-11-072-512-2541	Sequence 2541, Ap
13	1061.5	21.4	594	5	US-10-631-467-681	Sequence 681, App
14	1061.5	21.4	594	6	US-11-177-894-8	Sequence 8, Appli
15	1037.5	21.0	782	4	US-10-066-543-1424	Sequence 1424, Ap
16	912.5	18.4	475	4	US-10-104-047-3116	Sequence 3116, Ap
17	912.5	18.4	475	6	US-11-072-512-3116	Sequence 3116, Ap
18	905	18.3	642	4	US-10-108-260A-4483	Sequence 4483, Ap
19	905	18.3	642	6	US-11-177-894-9	Sequence 9, Appli
20	819.5	16.6	443	4	US-10-264-049-2917	Sequence 2917, Ap
21	817.5	16.5	179	5	US-10-495-663-1	Sequence 1, Appli
22	784.5	15.8	390	4	US-10-264-237-2758	Sequence 2758, Ap
23	735	14.8	139	3	US-09-957-708-31	Sequence 31, Appl
24	735	14.8	139	6	US-11-230-251-31	Sequence 31, Appl
25	684.5	13.8	483	4	US-10-108-260A-3990	Sequence 3990, Ap
26	594.5	12.0	660	4	US-10-108-260A-3644	Sequence 3644, Ap
27	573.5	11.6	500	5	US-10-450-763-35862	Sequence 35862, A
28	556	11.2	1604	3	US-09-746-491-8	Sequence 8, Appli
29	550	11.1	1338	6	US-11-097-143-37875	Sequence 37875, A
30	464.5	9.4	240	3	US-09-800-095A-38	Sequence 38, Appl
31	455.5	9.2	257	3	US-09-789-561-99	Sequence 99, Appl
32	455.5	9.2	257	3	US-09-833-245-2111	Sequence 2111, Ap
33	455.5	9.2	257	5	US-10-883-936-99	Sequence 99, Appl
34	455.5	9.2	257	6	US-11-264-096-2111	Sequence 2111, Ap
35	454	9.2	734	4	US-10-425-114-65967	Sequence 65967, A
36	438	8.8	646	4	US-10-189-718-6	Sequence 6, Appli
37	438	8.8	646	4	US-10-189-718-8	Sequence 8, Appli
38	438	8.8	646	6	US-11-097-143-20586	Sequence 20586, A
39	437	8.8	646	4	US-10-189-718-4	Sequence 4, Appli
40	418	8.4	378	5	US-10-821-273-58	Sequence 58, Appl
41	411.5	8.3	393	3	US-09-731-872-457	Sequence 457, App
42	411.5	8.3	393	3	US-09-876-997-457	Sequence 457, App
43	411.5	8.3	393	5	US-10-643-836-457	Sequence 457, App
44	397	8.0	912	4	US-10-408-765A-2280	Sequence 2280, Ap
45	394	8.0	640	4	US-10-437-963-157713	Sequence 157713,

ALIGNMENTS

RESULT 1

US-10-450-763-45847
 ; Sequence 45847, Application US/10450763
 ; Publication No. US20050196754A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hyseq, Inc
 ; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
 ; FILE REFERENCE: 790CIP3/US
 ; CURRENT APPLICATION NUMBER: US/10/450,763
 ; CURRENT FILING DATE: 2003-06-11
 ; PRIOR APPLICATION NUMBER: PCT/US01/08631
 ; PRIOR FILING DATE: 2001-03-30
 ; PRIOR APPLICATION NUMBER: 09/540,217
 ; PRIOR FILING DATE: 2000-03-31
 ; PRIOR APPLICATION NUMBER: 09/649,167
 ; PRIOR FILING DATE: 2000-08-23
 ; NUMBER OF SEQ ID NOS: 60736
 ; SOFTWARE: Custom
 ; SEQ ID NO 45847
 ; LENGTH: 898
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-450-763-45847

Query Match 75.5%; Score 3736; DB 5; Length 898;
 Best Local Similarity 82.3%; Pred. No. 0;
 Matches 727; Conservative 4; Mismatches 16; Indels 136; Gaps 6;

```

Qy      1 MRMAATAWAGLQGPPLPTLCPAVRTGLYCRDQAHAE----- 37
          |||||
Db      1 MRMAATAWAGLQGPPLPTLCPAVRTGLYCRDQAHAEATDVVLLAPFCQPKTRSHGTCPP 60

Qy      38 -----W---AMTSETS-----SG 47
          |  |  |
Db      61 TERDPRGEGSTEYPGRVDGIQGWGTRALTGWTDRLLCQACQTLPPRHWFLPGARGWLGG 120

Qy      48 SHCA-----RSRMLRRAQEEDSTVLIDVSPPEAEKRGSYGSTAH 87
  
```



```

; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHIKO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: 084335-0191
; CURRENT APPLICATION NUMBER: US/11/072,512
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: US 60/350,978
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: JP 2001-379298
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2574
; LENGTH: 920
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-072-512-2574

```

```

Query Match          30.9%; Score 1531.5; DB 6; Length 920;
Best Local Similarity 37.9%; Pred. No. 9.1e-139;
Matches 360; Conservative 168; Mismatches 316; Indels 105; Gaps 29;

```

```

Qy      44 TSSGSHCARSMLRRRAQEEDSTVLID----VSPPEAE-----KRGSYGST---AHASEP 91
      :|||  ::::  | :|  :| |  || |  :  | :  ||  |  |
Db      4 SSSGITNGKTKVFPVA--KDVNIFDELEAVSSPCKDDDSLHPGNLTSTSDASRLEA 61

Qy     92 GGQQAACRAGS-----PAKPRIADFLVWVEEDLKLDQRQDQAARDRTDMHRTWRETFD 146
      ||:  :  |  || |::||:  :  :  :|:  || |
Db     62 GGETVPERNKSNGLYFRDGKCRI-DYILVYRK-----SNPQTEK----REVFER 105

Qy    147 NLRAAGLCVDQDQVDQDNTTVHYALLSASWAVLCYYAEDLRLKLPQE----LPNQASNW 202
      |:|| || ::::  | : :  | | | || | || : ::| :  || :
Db    106 NIRAEGLMQMEKESLI-NSDIIFVKLHAPWEVLGRYAEQMNVRMPFRRKIYYLPRRYKFM 164

Qy    203 S-----AGLLAWLGIPNVLL--EVVPDVP--EYYSRFRVKNLPRFLGSDNQDTFFTST 253
      | : :  | |  | :  | :||:  : | :  | :  | :||| :
Db    165 SRIDKQISRLRRWLPKKPMRLDKETLPDLEENDCYTAPFSQQRHHFI-IHNKETFFNNA 223

Qy    254 KRHQILFEILAKTPYGHEKKNLLGIHQLLAEGVLSAAFPLHDGPFKTPPEGQAPRLNQR 313
      | : :  || :  | | | :||: || |  |||||:| : :  | |
Db    224 TRSRIVHHILQRIKY-EEGKNKIGLNRLLTNGSYEAAFPLHEGSYRSKNSIRTHGAENHR 282

Qy    314 QVLFQHWARWGKWNKYQPLDHVRRYFGEKVALYFAWLGFTGWLLPAAVVGTLVFLVGCF 373
      |: :  || | ||||| |||||: |||||: || | || | : | || |
Db    283 HLLYECWASWGVWYKYQPLDLVRRYFGEKIGLYFAWLGWYTGMLFPAAFIGLFVFLYGV 342

Qy    374 LVFSDIPTQELCGSKDSFEMCPLC-LDCPFWLLSSACALAQAGRLFDHGGTVFFSLFMAL 432
      :  ::::| : |  ||:|  || | || : | :  ||:| ||||: |||:
Db    343 TLDHSQVSKEVCQATDII-MCPVCDKYCPFMRLSDSCVYAKVTHLFDNGATVFAVMAV 401

Qy    433 WAVLLELYWKRKSATLAYRWDCSDYEDTEERPRPQFAAS-APMTAPNPITGEDEPYFPER 491
      || : ||: |||: | :|| | | :|: ||  ||| | :  ||:|: |||
Db    402 WATVFLEFWKRRRAVIAYDWDLIDWEEEEEEIRPQFEAKYSKKERMNPISGKPEPYQAF 461

Qy    492 SRARRMLAGSVVIVVMVAVVMCLVSIILYRAIMAVVSRSGNTLLA-AWA-----SRIA 545
      :  |::  :  | | :||:  : |::| :  : | | ||  |::|
Db    462 DKCSRLIVSASGIFFMICVVIAAVFGIVYRVVTV-----STFAAFKVALIRNNSQVA 514

Qy    546 SLTGSVV--NLVFILILSKIYVSLAHVLTRWEMHRTQTKFEDAFTLKVFIFQFVNFYSSP 603
      : ||: | |  |::|: :|  :| :|| |  ||:::|: |||:|: ||| |
Db    515 T-TGTAVCINFCIIMLLNVLYEKVALLLTNLEQPRTESEWENSFTLKMFLFQFVNLNSST 573

Qy    604 VYIAFFKGRFVGYPGNYHTLFG-VRNEECAAGGCLIELAQELLVIMVGKQVINNMQEVLI 662
      ||||| || | :|| | |  | || |  |||:|  : : ||| ||  | | :
Db    574 FYIAFFLGRFTGHPGAYLRLINRWRLEECHPSGCLIDLQMGMIMVLKQTNWNNFMELGY 633

```

```

Qy      663 PKLKGWQKFLRSKKRKAGASAGASQGPWEDDYELVPCE--GLFDEYLEMVLQFGFVTI 720
      | :: || : | :: | | || || | |||||:|||| |
Db      634 PLIQNWWR---RKVRQEHGPERKISFPQWEKDYNLQPMNAYGLFDEYLEMILQFGFTTI 690

Qy      721 FVAACPLAPLALLNNWVEIRLDARKEVCEYRRPVAERAQDIGIWFHILAGLTHLAVISN 780
      |||| |||| |||| :|||| | | ::||: | :||||: | | : |::|
Db      691 FVAAPLAPLALLNNIIEIRLDAYKFVTQWRRPLASRAKDIGIWFYGILEGIGILSVITN 750

Qy      781 AFLFASFSSDFLPRAYYRW-----TRAHDLRGFLNFTLA-----RAP 816
      ||::| :||:| | : : ||: | :
Db      751 AFVIAITSDFIPLRVYAYKYGPCAGQGEAGQKCMVGYVNASLSVFRISDFENRSEPESDG 810

Qy      817 SSFAAAHNRTCRYRAFRDDDGH----YSQTYWLLAIRLAFVIVFEHVFSVGRLLDLL 871
      | | : : ||| :| | : :|| | |||:||||:| | : | : |
Db      811 SEFSGTPLYKCRYRDYRPPHSLVPYGYTLQFWHVLAAALAFIIVFEHLVFCIKHLISYL 870

Qy      872 VPDIPESVEIKVKREYYLAKQALAENEVLFGTNGTKDEQPKGSELSSHW 920
      :||:| : : ::|| | | : : | | : : | : |
Db      871 IPDLPKDLDRMRREKYLIQEMMYEALERLQKERKERKKNKAHNEW 919

```

RESULT 4

US-11-177-894-7

```

; Sequence 7, Application US/11177894
; Publication No. US20060040292A1
; GENERAL INFORMATION:
; APPLICANT: West, et al.
; TITLE OF INVENTION: Tumor Markers and Uses Thereof
; FILE REFERENCE: 2002850-0048
; CURRENT APPLICATION NUMBER: US/11/177,894
; CURRENT FILING DATE: 2005-07-08
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7
; LENGTH: 960
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Transmembrane protein
US-11-177-894-7

```

```

Query Match          30.1%; Score 1488; DB 6; Length 960;
Best Local Similarity 37.6%; Pred. No. 1.6e-134;
Matches 363; Conservative 160; Mismatches 307; Indels 136; Gaps 28;

```

```

Qy      26 GLYCRDQAHAERWAMT--SETSSGSHCARSRMLRRRAQEEDSTVLIDVSPPEAEKRGSYG 83
      ||| || : : : || :| || | | : |
Db      52 GLYFRDGRRKVDYILVYHHKRPSG-----NRTLVRVQHS DTP-----SGA 92

Qy      84 STAHASEPGGQQAACRAGSPAKPRIADFLVWEECLKLDRQQDQAARDRTDMHRTWRET 143
      : | : | : || | :| :| | ||
Db      93 RSVKQDHPPLPGKGASLDAGSGEPP-----MDYHEDD-----KRFRREE 130

Qy      144 FLDNLRAAGLCVDQDQDVGNTTVH---YALLSASWAVLCYYAEDLRLKLPLQELPNQAS 200
      : || ||| :: :| :| :| : : | || | | | :| :| : :: :
Db      131 YEGNLEAGLELE---RDEDTKIHGVGFKIHAPWNVLCREAEFLKLMPTKKMYH--I 184

Qy      201 NWSAGLLAWLGIPNVLLEVVPDVPPEYYSCR-----FRVNKLPRFLGSDNQDTFF 250
      | : ||| | :|| : : : | : | | | | :| :|
Db      185 NETRGLLK--KINSVLQKITDPIQPKVAEHRPQTMKRLSYFPSREKQHLFDLSD-KDSFF 241

Qy      251 TSTKRHQILFEILAKTPYGHEKKNLLGIHQLLAEGVLSAAFPLHDGPFKTPEGPQAPRL 310
      | | | :||| :| | : :|| || | :||:|||| : :
Db      242 DSKTRSTIVYEILKRTTCTKAKYS-MGITSLLANGVYAAAYPLHDGDY-----NGENVEF 295

Qy      311 NQRQVLQFQHWARWGKWNKYQPLDHVRRYFGEKVALYFAWLGFYTGWLLPAAVVGTLLVFLV 370
      | :| :| : ||| : | ||| : | :||| : ||||| | | :| :| :| :||
Db      296 NDRKLLYEEWARYGVFYKYQPIDLVRKYFGEKIGLYFAWLGVTQMLIPASIVGIIIVFLY 355

Qy      371 GCFLVFSDIPTQELCGSKDSFEMCPLC-LDCPFWLLSSACALAQAGRLFDHGGTVFFSLF 429
      || : :|| :| : : |||| | :| :||| | :| ||| : ||||:|
Db      356 GCATMDENIPMEMCDQRHNITMCPLCDKTC SYWKMSACATARASHLFDNPATVFFSVF 415

Qy      430 MALWAVLLEWYKRKSATLAYRWDCSDYEDTEE----RPRPQFAA-----SAPMTAPNPI 480
      ||||| :||| | ||| : :| | || : : | | : |
Db      416 MALWAATFMEHWKRKQMRNLNRYRWDLTGFEVEEAVKDHPRAEYEAARVLEKSLKKESRNKE 475

```



```
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15228
; LENGTH: 1219
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-15228
```

```
Query Match      29.2%; Score 1445; DB 6; Length 1219;
Best Local Similarity 35.6%; Pred. No. 3.6e-130;
Matches 342; Conservative 165; Mismatches 332; Indels 122; Gaps 27;
```

```
Qy      35 AERWAMTSETSSGSHCARSRML-----RRRAQEEDSTVLIDVSPPEAEKRGSY 82
      |:| :| || |:| |:| :| |:| |||
Db      249 ADRVNQSYEVMESSH---SNVLPDQFGYRQLIPTERKASDTASSV-----SGSY 294

Qy      83 GSTAHASEP---GGQAAACRAGSPAKP-----RIADFVLVW-EEDLKLDRQ 125
      : ||: ||: | : | || | ||| : :
Db      295 YGSRKASKSNSLGGESGDERRVSKQDREGLDPESLMFRDGRRKVDMLAWEEEDLGVMT 354

Qy      126 QDAAARDRTDMHRTWRETFLDNLRAAGLCVDQDD-VQDGNNTTVHYALLSASWAVLCYYAE 184
      :: || |:|:| || |:| | | : : | : ||
Db      355 AEAKRRDN-----RRSFMENLIKEGLEVELEDKQSFSNEKTFFLKIHLPWRLETRLAE 407

Qy      185 DLRLKLP-----LQELPNQASNWSAGLLAWLGIPNVLLVVPDVP 225
      : |||| |:| | | : : |||
Db      408 VMNLKLPVKRFITISVKPSWDEENVVLRNMQYWKDVWQR-LTKKIQLDQTLLE---GET 462

Qy      226 EYYSRCFRVNKLPRFLGSDNQDFTFTSTKRHQILFEILAKTPYGHEKKNLLGIHQLLAEG 285
      : : | :|: | :| ||| :| : : :| :||: : : || :|: :|
Db      463 TFKAATANGNPEEQFIVKD-RATAFTSAQRSLMVMQVLIRTPFDESQRS--GIRRLMNDG 519

Qy      286 VLSAAFPLHDGPFKTPPEGPQAPRLN-QRQVLFQHWARWGKWNKYQPLDHVRRYFGEKVA 344
      ||||:| : : | : : : |:|:| || :| | || ||:|:|:|:|
Db      520 TYLGCFFLHEGRY---DRPHSSGISLDRRVLYQTWAHPSQWYKQPLCLVRKYFGDKIA 575

Qy      345 LYFAWLG FYTGWLLPAAVVGTLVFLVGCFLVFS--IPTQELCG--SKDSFEMCPLC-LD 399
      ||| ||||| |:| ||||| |:| : |:| |:|:| : :|||
Db      576 LYFCWLGFYTEMLVYPVAVVGTLCFIYGLATLESDNTPSKEICNEYGTGNITLCPLCDKA 635

Qy      400 CPFWLLSSACALAQAGRLFDHGTVFFSLFMALWAVLLLEYWKRKSATLAYRWDCSDYED 459
      | : || :| : : ||| :|:|:|:| || |||| : | : | : |
Db      636 CSYQRLSESCLSRLTYLFDNPSTVFFAIFMSFWATTFLLEWKRKQSVLVWEWDLHNV-D 694

Qy      460 TEERPRPQFAASAPMTAPNPITGEDEPYFPERSRARRMLAGSVVIVVMVAVVMCLVSII 519
      :| ||:| :| ||| :|:| : : :|:|:| : : |
Db      695 MDEENRPEFETNATTFRMPVTREKEPYMSTWNRSIRFVITGSAVLFMISVVL SAVLGTI 754

Qy      520 LYRAIMAIIVVSRSGNTLLAAWASRIASLTGSVNVLVFILILSKIYVSLAHVLTRWEMHRT 579
      ||| : | : : ||| :| :||| |:|:|:| :| || | ||
Db      755 LYRITLVSVIYGGGGFFVKEHAKLFTSVTAALINLVVIMILTRIYHRMAIKLTNLENPRT 814

Qy      580 QTKFEDAFTLKVFI FQFVN FYSSPVYIAFFKGRFVGYPGNYHT---LFGVRNEECAAGG 635
      |:|:|:| |:| |:|:|:| :|:|:|:| ||| :|:|:| :|
Db      815 HTEYEDSYTFKIFFEFMNFYSSLIYIAFFKGRFFDYPGDDQARKSEFFRLKNDICDPAG 874

Qy      636 CLIELAQELLVIMVGKQVINNMQEVLPKLGWQKFRRLRSKKRKAGASAGASQGPWEDD 695
      || || :| :|:|:| || | || || : :| :| :|
Db      875 CLSELQICLAIIMVGKQCWNFMFYLFPKFWNWR---QRKHQATKDESHLMAWEQD 930

Qy      696 YELV-PCE-GLFDEYLEMVLQFGFVTIFVAAACPLAPL FALLNNWVEIRLDARKFVCEYRR 753
      | : | |||||:|:|:|:| ||||| ||||| | | : ||
Db      931 YHMQDPGR LALFDEYLEMILQYGFVTLFVAAFPLAPL FALLNNVAEIRLDAYKMVTQARR 990

Qy      754 PVAERAQDIGIWFHILAGLTHLAVISNAFLAFSSDFLPRAYYR--WTRAHDLRGFLNFT 811
      |:|:| :|| |:| |:| |:|:|:|:|:|:| |:| : : | | : :
Db      991 PLAERVEDIGAWYGILRIITYTAVVSNAFVIAYTSDFIPRMVYKFVYSETHLAGYIEHS 1050

Qy      812 LA-----RAPSSFAAHNRTCRYRAFRDDDGHY-----SQTYNLLAIRLAFVIVF 857
      |:| :| :| :|:| :|:| | | :|:| :|:|:|
Db      1051 LSIFNTSDYKEEWGASVSEKDPDTCQYRGYRNGPKDYEPYGLSPHYWHVFAARLAFVVVF 1110

Qy      858 EHVVSFVGRLLDLLVPDIPESVEIKVKREYLLAKQALAENEVLFGTNGTKDEQPKGSELS 917
      |||| : : :|:| |:| :||| |||:| : :| | | : :
Db      1111 EHVVFVITGIMQFIIPDVPSEVKTQMQRQLLAKEAKYQ-----HGKRAQGDSQDIM 1163
```


Qy 918 S 918
|
Db 1164 S 1164

RESULT 7

US-10-484-148-14

; Sequence 14, Application US/10484148

; Publication No. US20040248251A1

; GENERAL INFORMATION:

; APPLICANT: LAL, Preeti G.; HONCHELL, Cynthia D.;

; APPLICANT: FORSYTHE, Ian J.; CHAWLA, Narinder K.;

; APPLICANT: TANG, Y. Tom; BOROWSKY, Mark L.; BARROSO, Ines;

; APPLICANT: YUE, Henry; WARREN, Bridget A.;

; APPLICANT: THANGAVELU, Kavitha; GIETZEN, Kimberly J.;

; APPLICANT: AZIMZAI, Yalda; LEE, Ernestine A.;

; APPLICANT: BAUGHN, Mariah R.; GORVAD, Ann E.;

; APPLICANT: DUGGAN, Brendan M.; TRAN, Bao;

; APPLICANT: LI, Joana X.; RICHARDSON, Thomas W.;

; APPLICANT: ELLIOTT, Vicki S.; ZEBARJADIAN, Yeganeh

; APPLICANT: TRAN, Uyen K.; YAO, Monique G.;

; APPLICANT: PETERSON, David P.; LUO, Wen

; APPLICANT: LEHR-MASON, Patricia M.

; TITLE OF INVENTION: RECEPTORS AND MEMBRANE ASSOCIATED PROTEINS

; FILE REFERENCE: PF-1082 USN

; CURRENT APPLICATION NUMBER: US/10/484,148

; CURRENT FILING DATE: 2004-01-15

; PRIOR APPLICATION NUMBER: PCT/US02/22833

; PRIOR FILING DATE: 2002-07-16

; PRIOR APPLICATION NUMBER: US 60/306,020

; PRIOR FILING DATE: 2001-07-17

; PRIOR APPLICATION NUMBER: US 60/308,179

; PRIOR FILING DATE: 2001-07-27

; PRIOR APPLICATION NUMBER: US 60/309,702

; PRIOR FILING DATE: 2001-08-02

; PRIOR APPLICATION NUMBER: US 60/311,476

; PRIOR FILING DATE: 2001-08-10

; PRIOR APPLICATION NUMBER: US 60/311,718

; PRIOR FILING DATE: 2001-08-10

; PRIOR APPLICATION NUMBER: US 60/311,551

; PRIOR FILING DATE: 2001-08-10

; PRIOR APPLICATION NUMBER: US 60/314,798

; PRIOR FILING DATE: 2001-08-24

; PRIOR APPLICATION NUMBER: US 60/316,639

; PRIOR FILING DATE: 2001-08-31

; PRIOR APPLICATION NUMBER: US 60/317,996

; PRIOR FILING DATE: 2001-09-07

; NUMBER OF SEQ ID NOS: 46

; SOFTWARE: PERL Program

; SEQ ID NO 14

; LENGTH: 910

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc_feature

; OTHER INFORMATION: Incyte ID No: 3718011CD1

US-10-484-148-14

Query Match 28.3%; Score 1402.5; DB 5; Length 910;

Best Local Similarity 38.2%; Pred. No. 3.1e-126;

Matches 322; Conservative 157; Mismatches 286; Indels 79; Gaps 24;

Qy 108 RIADFVLVWEEDLKLDRQQDSAAARDRTDMHRTWRETFLDNLRAAGLCVD-QQDVQDGNNT 166

Db 67 RRIDFVLVYED----ESRKETNKKGTNEKQRRKRQAYESNLICHGLQLEATRSVLDDKLV 122

Qy 167 VHYALLSASWAVLCYYAEDLRKLPLQELPNQASNWSA--GLLAWLGI PNVLLEVVPDVP 224

Db 123 --FVKVHAPWEVLCTYAEIMHIKLPK--PNDLKNRSSAFGTLNWF TKVLSVDESIIKPE 178

Qy 225 PEYYSRFRVNKLPRFLGSDNQDTFFTSTKRHQILFEILAKTPYG-HEKKNLLGIHQLLA 283

Db 179 QEFTTAPFEKRNMDFYIVD-RDAFFNPATRSRIVYFILSRVKYQVINNVSKFGINRLVN 237

Qy 284 EGVLSAAFLHDGPFKTPPEGPAAPRLNQRQVLFQHWARWGKWNKYQPLDHVRRYFGEKV 343

Db 284 EGVLSAAFLHDGPFKTPPEGPAAPRLNQRQVLFQHWARWGKWNKYQPLDHVRRYFGEKV 343

; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 24771
; LENGTH: 1075
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-24771

Query Match 27.7%; Score 1369.5; DB 6; Length 1075;
Best Local Similarity 37.4%; Pred. No. 6.4e-123;
Matches 313; Conservative 163; Mismatches 283; Indels 77; Gaps 20;

```
Qy      108 RIADFLVLWEEDLKLDRQQDQAARDRTDMHRT-WRETFLDNLRAAGLCVD--QQDVQDGN 164
      | ||| : : | : | : | | | : || : :|
Db      192 RSIDFVLAYRIN-----AHEPTELENTEKRRVFANLISQGLEVESSQKD----- 236

Qy      165 TTVHYALLSASWAVLCYYAEDLRKLPLQELPNQASNWSAGLLAWLGIPNVL-----LE 218
      : : : | || ||| |::|::|:| : : : :| :
Db      237 -QIWVFKIHAPLEVLRRYAEILKLRMPMKEIPGMSVVRSTKSVFSSLKHVFQFFLRNIY 295

Qy      219 VVPDPVPEYYSCRFRV--NKLPRFLGSDNQDTFFTSTKRHQILFEIL--AKTPYGHEKKN 274
      | ::|: : || : : | |||: | :| || :| ::
Db      296 VDEEIFPK-RAHRFTAIYSRDKEYLFDIRQDCFFTAVRSRIVEFILDQRQFPAKNQHDM 354

Qy      275 LLGIHQLLAEGVLSAAFLPHDGPFTKPEGPQAPRLNQRQVLFQHWARWGKWNKYQPLDH 334
      || :|:|||| ||:|||| | :|:|||| || :||||
Db      355 AFGIERLIAEGVLSAAFLPHDGEITETG-----TMRALLYKHVASVPKWRYQPLDD 406

Qy      335 VRRYFGKEVALYFAWLGFYTGWLLPAAVVGTLVFLVGCFLVFSDIPTQELCGS-KDSFEM 393
      :: ||| | : |||||:| || |::| : || | : : :| :|| : |
Db      407 IKEYFGVKIGLYFAWLGYYTYMLLLASIVGVICFLYSWFSLKNYVPVKDICQSGNTNITM 466

Qy      394 CPLCLDCPFWLLSSACALAQAGRLFDHGGTVFFSLFMALWAVLLLEYWKRKSATLAYRWD 453
      |||| | || | | : | : | : |||:|: || | || ||| || : :||
Db      467 CPLCDWCNFWDLKETCNYAKVTYLIDNPSTVFFAVFMSFWATLFLELWKRYSAEITHRWD 526

Qy      454 CSDYEDTEERPRQFAA---SAPMTAPNPITGEDEPYFP-ERSRARRMLAGSVVIVMVA 509
      : :: || |||: | | : :| || | | : : :| ::::|
Db      527 LTGFVDVHEHPRQYLARLEHIPTRVDYVNIKEPTVPFWRMKLPATVFSFSVLLLIA 586

Qy      510 VVMCLVSIILYRAIMAVVSRSGNTLLAAWASRIASLTGSVNVLFILILSKIYVSLAH 569
      : : |:::| | : : : : | :|: : : || : ||: | ||
Db      587 LAFVALLAVVYRMSMLAALKVGASPMITSSAIVLATASAAFVNLCLLYILNYMNLAE 646

Qy      570 VLTRWEMHRTQTKFEDAFTLKVFIFQVNFYSSPVYIAFFKGRFVGYPGNYHTLFGVRNE 629
      || || |||:|:| :|:| :|:|:| :|:|:|:|:|:| | :| | |
Db      647 YLTELEMWRTQTFDDSLTLKIYLLQFVNYIASIFYIAFFKGKFVGHGPEYNKLFDYRQE 706

Qy      630 ECAAGGCLIELAQELLVIMVGKQVINNMQEVLIPLKLGWWQKFLRLSKKRKAGASAGASQ 689
      |:|:|:| : :|:|:| :|:|:| :| :| :|:| :|:|:| :| :|
Db      707 ECSSGGCLTELICQLAIIMVGKQAFNTILEVYLP---FWRKV---LAIQVGLSRLFNN 759

Qy      690 GP-----WEDDYELVP--CEGLFDEYLEMVLQFGFVTIFVAACPLAPLALLNNV 738
      | :|:| : || |||||:|:|:|:|:|:|:|:|:|:|:|:| :
Db      760 TPNPDKAKDERWMRDFKLLDWGARGLFPEYLEMVLQYGFVTIFVAAFPLAPFFALLNNIL 819

Qy      739 EIRLDARKEVCEYRRPVAERAQDIGIWFHILAGLTHLAVISNAFLAFSSDFLPRAYRW 798
      |:|:|:| : :|:|:| :|:|:| :| :|:|:| :|:|:| :| ||
Db      820 EMRLDAKKLLTHHKRPVSQRVRDIGVYRILDCIGKLSVITNGFIIAFTSDMIPRLVYRH 879

Qy      799 --TRAHDLRGFLNFTLAR-----APSSFAAAHN----RTCRYRAFR-----DDGHYSQT 842
      : | :|:|:| :|:|:| :|:|:| :| :| :| :| :|
Db      880 YVNKQGTLDGYLNFTLSEFKVIDSPTLYSLAGDLSNITTCRYTDFRLPSSPEKYTLSSM 939

Qy      843 YWNLLAIRLAFVIVFEHVVSFVGRLLDLLVPDIPESVEIKVKREYYLAKQALAENE 898
      : : || || ||:|:| | | :|:| : : :| :| : : :|
Db      940 FYIILACRLGFVVVFENFVALVMILVRWCIPDMSVELRDQIRREVYVTNEIIDQE 995
```

RESULT 9
US-11-177-894-10
; Sequence 10, Application US/11177894
; Publication No. US20060040292A1
; GENERAL INFORMATION:
; APPLICANT: West, et al.
; TITLE OF INVENTION: Tumor Markers and Uses Thereof

```
; FILE REFERENCE: 2002850-0048
; CURRENT APPLICATION NUMBER: US/11/177,894
; CURRENT FILING DATE: 2005-07-08
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 10
; LENGTH: 712
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Homo sapiens
US-11-177-894-10
```

```
Query Match      27.6%; Score 1367.5; DB 6; Length 712;
Best Local Similarity 41.6%; Pred. No. 5.4e-123;
Matches 299; Conservative 128; Mismatches 220; Indels 71; Gaps 17;
```

```
Qy      259 LFEILAKTPYGHEKKNLLGIHQLLAEGVLSAAFPLHDGPFKTPPEGQPAPRLNQRQVLFQ 318
      ::||| :| | : :|| || | :||:|||| : : | :|:|:
Db      2 VYEILKRTTCTKAKYS-MGITSLLANGVYAAAYPLHDGDY-----NGENVEFNDRKLLYE 55

Qy      319 HWARWGKWNKYQPLDHVRRYFGEKVALYFAWLGFYTGWLLPAAVVGTLVFLVGCFLVFS 378
      |||:| : |||:| ||:||||: ||||| || |:|:|:| :||| || : :
Db      56 EWARYGVFYKYQPIDLVRKYFGEKIGLYFAWLGVTQMLIPASIVGIIIVFLYGCATMDEN 115

Qy      379 IPTQELCGSKDSFEMCPLC-LDCPFWLLSSACALAQAGRLFDHGGTVFFSLFMALWAVLL 437
      ||: |:| : : |||| | :| :|||| |:| ||: ||||:|||||
Db      116 IPSMEMCDQRHNITMCPKCDKTCYSWKMSACATARASHLFDNPATVFFSVFMALWAATF 175

Qy      438 LEYWKRSATLAYRWDCSDYEDTEE----RPRPQFAA-----SAPMTAPNPITGEDEPYF 488
      :|:|||| | ||| : :|: || || :| | : | | :|
Db      176 MEHWKRKQMLNRYRWDLTGFEIEEEAVKDHPRAEYEARVLEKSLKESRNKET--DKVKL 233

Qy      489 PERSRARRMLAGSVVIVVMVAVVMCLVSIILYRAIMAVVSRSGNTLLAAWASRIASLT 548
      | | | | :|: || : :|: || : :|: || : :|: || : :|
Db      234 TWRDRFPAYLTNLVSIIFMIAVTFAIVLGVIIYRISMAAALAMNSSPSVRSNIRVTVTAT 293

Qy      549 GSVVNLVFIILSKIYVSLAHVLTWRMHRTQTKFEDAFTLKVFIFQFVNFYSSPVYIAF 608
      :||| |:| :| :| | :|: || : | | :||| | : | :|
Db      294 AVIINLVVILLDEVYGCIARWLTKEIEVPKTEKSFEERLIFKAFLLKFVNSYTPIFYVAF 353

Qy      609 FKGRFVGYPGNYHTLF-GVRNEECAAGGCLIELAQELLVIMVGKQVI-NMQEVLIPKLK 666
      ||||| ||| :| | ||| ||||:| :| :||:|:| ||: | : ||:|
Db      354 FKGRFVGRPGDYVYIFRSFRMEECAPGGCLMELCQLSIIMLGKQLIQNNLFEIGIPKMK 413

Qy      667 GWWQKFLRLSKRKAGASAGASQGPWEDDYELVPCEGLFDEYLEMVLQFGFVTIFVAACP 726
      : :|: : : :| || | | || ||:|:|:|:|:|:|:|
Db      414 KLIRYLKQKQSPDHEECVKKRQRYEVDYNLEPFAGLTPEYMEMIIQFGFVTLFVASFP 473

Qy      727 LAPLFALLNNWVEIRLDARKFVCEYRRPVAERAQDIGIWFHILAGLTHLAVISNAFLAF 786
      ||||| ||| :|||:| ||| |||:|:|:| ||| |||:|
Db      474 LAPLFALLNNIIEIRLDAKKFVTELRRPVAVRAKDIGIWNILRGIGKLAVIINAFVISF 533

Qy      787 SSDFLPRA--YYRWTRAHDLRGFLNFTLARAPSSF-----AAAHN-----RTCR 828
      :|||:| | :| : | :|: || || | | :|
Db      534 TSDFIPRLVYLYMSKNGTMHGFVNHTL----SSFNVSDFQNGTAPNDPLDLGYEVQICR 589

Qy      829 YRAFRD---DDGHY--SQTYWNLAIRLAFVIVFEHVVSFVGRLLDLLVPDIPESVEIKV 883
      |: |:| : | |: :| :| |||||:|:| : :| :|||: : :
Db      590 YKDYREPPWSENKYDISKDFWAVLAARLAFVIVFQNLVFMMSDFVDWVIPDIPKDISQEI 649

Qy      884 KREYYLA-----KQALAENEVLFGTNGTKDEQP-----KGSELSSH 919
      :| | | || | | ||| | || |
Db      650 HKEKVLVVELFMREEQDKQQLL--ETWMEKERQKDEPPCNHNTKACPSLGSFAPSH 705
```

RESULT 10

US-11-097-143-21858

; Sequence 21858, Application US/11097143

; Publication No. US20050208558A1

; GENERAL INFORMATION:

; APPLICANT: Venter, J. Craig

; APPLICANT: et al.

; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID

; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE

; TITLE OF INVENTION: DROSOPHILA GENES.

```
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 21858
; LENGTH: 1058
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-21858
```

```
Query Match      24.2%; Score 1199.5; DB 6; Length 1058;
Best Local Similarity 33.6%; Pred. No. 2.1e-106;
Matches 294; Conservative 155; Mismatches 280; Indels 147; Gaps 26;
```

```
Qy      108 RIADFVLVWEEDLKLDRQQDSAARDRTDMHRTWRETFLDNLRAAGLCVDQDDVQDGNTTV 167
      | | | | : : : | | | | | : | | : : | | |
Db      204 RSVDFVLAYNGETQLEE-----HRRKCEIFEANLQREGLQLEHNKQV----RV 247

Qy      168 HYALLSASWAVLCYYAEDRLRLPLQELPNQASNWSAGLLAWLGIPNVLLEVVPDVPPEY 227
      | : : | | | | | : : | | : : | | : : | |
Db      248 HFIKIHAPAEVLYRYAEILKIKVPLKPIPGQD-----QIFAESAHEF 289

Qy      228 YSCRFRVNKL-----PRF-----LGSDNQDTFFTSTKRHQILFE 261
      : | | : | | | | | | | | | | | | : : | | : |
Db      290 KTCFSRMCKSLFSSVQLNTALFPEREPRIHLEFSRNYLELYDTEHPNFDASTRYSIINF 349

Qy      262 ILAKTPY--GHEKKNLLGIHQLLAEGVLSAAFPPLHDGPFKTPPEGPQAPRLNQ-----RQ 314
      | | : : | | : | | : | | : | | : | | : : | |
Db      350 ILQQRQFVEGEETADNLGIEKLVQDGVYTCAYTLHDVERRSRSAAGVGQHIQVEEQRLRE 409

Qy      315 VLFQHWARWGKWNKYQPLDHVRRYFGEKVALYFAWLGFYTGWLLPAAVVGTLVFLVGCFL 374
      | : | | | | : | | | | | | | | | | | | | |
Db      410 TLKPFYC-----SLQPLDQIKDYFGAKVALYFAWLGFYQMLPIISVFGVLCFLYGFIT 463

Qy      375 VFSDIPTQELCGSKDSFEMCPLC-LDCPFWLLSSACALAQAGRLFDHGGTVFFSLFMALW 433
      | | : : : | : | | | | : | | : : | | : | | : |
Db      464 WNSDPISRDI CDNGTI-MCPQCDCRS CDYWR LNETCTSSKFNYLIDNNMTVVFAFSMAIW 522

Qy      434 AVLLLEYWKRKSATLAYRWDCSDYEDTEERPRPQFAASAPMTAPNPITGEDEPYFPERSR 493
      | | : | | | | | : | | : : | | | | | : | : :
Db      523 AVVYLEFWKRYSAGLVHRWGTLTGFTTHVHRPQYLARISRT--KKLAG--KAYEQDHTG 578

Qy      494 ARRMLAGSV-----VIVVMVAVVVMCLVSIILYRAIMAVVSRSGNTLLA 538
      . | : | | | : : | : : | : : | | : | : : |
Db      579 KRTILDPPVPFWSFKFLPNFTSYSIMVLFICISVIAIAGIIIR-----MAQRASHSILG 633

Qy      539 AWAS-----RIASLTGSVVNLVFILILSKIYVSLAHVLTRWEMHRTQTKFEDAFTLKVFI 593
      : | : | : : | : | : : | : | : | : | : | : :
Db      634 SENSMTFKVMILPMTAGIIDLIVISLDMVYSNLAVKLTNYEYCRTQTEYDESLTIKNYV 693

Qy      594 FQFVNFYSSPVYIAFFKGRFVGYPGNYHTLFGVRNEECAAGGCLIELAQELLVIMVGKQV 653
      | | | | : | | | | | | | : : | | | | | | : : | |
Db      694 FQFVNYSSLFYIAFLKGKFGVYPAKYNRVLGFRQEECPGGCLMELCMQLVIIMAGKQA 753

Qy      654 INNMQEVLIPKL-----KGWW---QKFRRLRSKKRKAGASAGASQGPWEDDYELVP 700
      : | : | | | | | : : | | | | : : | | : : | |
Db      754 VNAIVEMLIPYLMRTFKELSYRHGWYKSHQDQRL-----VPYNQFTEDYNLLP 801

Qy      701 CE--GLFDEYLEMVLQFGFVTIFVAACPLAPLFA LLNNWVEIRLDARKFVCEYRRPVAER 758
```

```

      | : |||||:||||:| | |||| |||| :|:|||| | : |||| |
Db      802 AENNSLYVEYLEMNVQGFITLFSLAFPLAPLLALLNNVIEVRLDAIKMLRFLRRPVGMR 861

Qy      759 AQDIGIWFHILAGLTHLAVISNAFLAFSSDFLPRAYYRWTRAH-DLRGFLNFTLA---- 813
      |:||||| | : :| :|| :| :|||: :|: | : :| :|||||
Db      862 ARDIGVWSHIMTVVTRIAVASSAMIIAFSTNLIPKIVYAASMGDPENNYLNFTLAVFNT 921

Qy      814 ----RAPSSFAAAH-NRT-CRYRAFRD--DDGH-YSQ--TYWNLLAIRLAFVIVFEHVVF 862
      | : ||||| | : :| :|| :| :||| : || :| ||||: : : :
Db      922 KDFQVQPLLGGSQHVNETVCRYTEFRNSPEDPHYPKRPIMYWKILTGRLAFIVYQNIIT 981

Qy      863 SVGRLLDLLVPDIPESVEIKVKREYYLAKQALAENE 898
      : :| ||| : : ||| :| : : | |
Db      982 MLQGILRWAVPDVSGRLLKRIKRENFLLREHIIEYE 1017.

```

RESULT 11

```

US-10-104-047-2541
; Sequence 2541, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20030236392A1el full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2541
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-2541

```

```

Query Match      23.3%; Score 1154; DB 4; Length 596;
Best Local Similarity 41.3%; Pred. No. 2.3e-102;
Matches 250; Conservative 108; Mismatches 194; Indels 54; Gaps 14;

```

```

Qy      357 LLPAAVGTVLVFLVGCFLVFSDIPTQELCGSKDSFEMCPLC-LDCPFWLLSSACALAQAG 415
      | ||| :| ||| | : :||:| | |||:| ||| || :| | :
Db      2 LFPAAFIGLFLVFLYGVTTLDHSQVSKVCQATDII-MCPVCDKYCPFMRLSDSCVYAKVT 60

Qy      416 RLFDHGGTVFFSFLMALWAVLLELYWKRSATLAYRWDCSDYEDTEERPRPQFAAS-APM 474
      |||:| ||||:||||:| : ||:|||| | :|| || |:|: || |||| | :
Db      61 HLFDNATVFFAVFMAVWATVLEFWKRRRAVIAYDWDLIDWEEEEIEIRPQFEAKYSKK 120

Qy      475 TAPNPITGEDEPYFPERSRARRMLAGSVVIVVMVAVVMCLVSIILYRAIMAVVSRSGN 534
      |||:| ||| : ||: : | | : ||: : |: || :
Db      121 ERMNPISGKPEPYQAFDCKSRLIVSASGIFFMICVVIAAVFGIVYIRVTV-----S 173

Qy      535 TLLA-AWA-----SRIASLTGSVV--NLVFILILSKIYVSLAHVLTWRWEMHRTQTKFEDA 586
      | | | | ||| :||:| ||: | | ||:|: :| :| :|| | ||: :||:|
Db      174 TFAAFKWALIRNNSQVAT-TGTAVCINFCIIMLLNVLYEKVALLLTNLEQPRTESEWENS 232

Qy      587 FTLKVFIQFVNIFYSSPVYIAFFKGRFVGYPGNYHTLFG-VRNEECAGGCLIELAQELL 645
      ||||:| |||| || ||||| ||| :|| | | | ||| ||||:| :
Db      233 FTLKMFQLQFVNLSSTFYIAFFLGRFTGHPGAYLRLINRWRLLEECHPSGCLIDLQMOMG 292

Qy      646 VIMVGKQVINNMQEVLPKLGWQKFRRLRSKKRKAGASAGASQGPWEDDYELVPCE--G 703
      :||| || || | : : || : | : : | | || || || |
Db      293 IIMVLKQTWNNFMELGYPLIQNWTR---RKVRQEHGPERKISFPQWEKDYNLQPMNAYG 349

Qy      704 LFDEYLEMVLQFGFVTIFVAACTPLAPLALLNNVVEIRLDARKFVCEYRRPVAERAQDIG 763
      |||||:| |||| ||||| |||| |||| :||| || || :|||:| ||:|
Db      350 LFDEYLEMILQFGFTTIFVAAFPLAPLLALLNNIIEIRLDAYKFVTQWRRPLASRAKDIG 409

Qy      764 IWFHILAGLTHLAVISNAFLAFSSDFLPRAYYRW-----TRAHDLRGFLNFTLA 813
      ||: || | : ||:||||:| :||:| | : : ||:| :|
Db      410 IWYGILEGIGILSVITNAFVIAITSDFIPRLVYAYKYGPCAGQGEAGQKCMGVYNASLS 469

Qy      814 -----RAPSSFAAAHNRTCRYRAFRDDDGH-----YSQTYWNLLAIRLAFV 854
      | | : : |||| :|| | : :||:| ||||:
Db      470 VFRISDFENRSEPDGSEFSGTPLKYCRYRDRDPHSLVPYGYTLQFVHVLAAARLAFI 529

```

Qy 855 IVFEHVVS VGRLLDLLVPDIPESVEIKVKREYYLAKQALAENEVLFGTNGTKDEQPKGS 914
 |||||:| : | : |:|: : ::|| || : : | : | : : |
 Db 530 IVFEHLVFCIKHLISYLPDLPKDLRDRMRREKYLIQEMMYEALERLQKERKERKNGK 589

Qy 915 ELSSHW 920
 : |
 Db 590 AHHNEW 595

RESULT 12

US-11-072-512-2541
 ; Sequence 2541, Application US/11072512
 ; Publication No. US20060029945A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ISOGAI, TAKAO
 ; APPLICANT: SUGIYAMA, TOMOYASU
 ; APPLICANT: OTSUKI, TETSUJI
 ; APPLICANT: WAKAMATSU, AI
 ; APPLICANT: SATO, HIROYUKI
 ; APPLICANT: ISHII, SHIZUKO
 ; APPLICANT: YAMAMOTO, JUN-ICHI
 ; APPLICANT: ISONO, YUUKO
 ; APPLICANT: HIO, YURI
 ; APPLICANT: OTSUKA, KAORU
 ; APPLICANT: NAGAI, KEIICHI
 ; APPLICANT: IRIE, RYOTARO
 ; APPLICANT: TAMECHIKA, ICHIRO
 ; APPLICANT: SEKI, NAOHIKO
 ; APPLICANT: YOSHIKAWA, TSUTOMU
 ; APPLICANT: OTSUKA, MOTOYUKI
 ; APPLICANT: NAGAHARI, KENJI
 ; APPLICANT: MASUHO, YASUHIKO
 ; TITLE OF INVENTION: Novel full length cDNA
 ; FILE REFERENCE: 084335-0191
 ; CURRENT APPLICATION NUMBER: US/11/072,512
 ; CURRENT FILING DATE: 2005-03-07
 ; PRIOR APPLICATION NUMBER: US 60/350,978
 ; PRIOR FILING DATE: 2002-01-25
 ; PRIOR APPLICATION NUMBER: JP 2001-379298
 ; PRIOR FILING DATE: 2001-11-05
 ; NUMBER OF SEQ ID NOS: 4096
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2541
 ; LENGTH: 596
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-11-072-512-2541

Query Match 23.3%; Score 1154; DB 6; Length 596;
 Best Local Similarity 41.3%; Pred. No. 2.3e-102;
 Matches 250; Conservative 108; Mismatches 194; Indels 54; Gaps 14;

Qy 357 LLPAAVVGLTVFLVGCFLVFSDIPTQELCGSKDSFEMCPLC-LDCPFWLLSSACALAQAG 415
 | |||:| ||| : : |:|: | |||:| ||| ||:| | :
 Db 2 LFPAAFIGLGFVFLYGVTTLDSQVSKEVCQATDII-MCPVCDKYCPFMRLSDSCVYAKVT 60

Qy 416 RLFDHGGTVFFSLFMALWAVLLEIYWKRSATLAYRWDCSDYEDTEERPRPQFAAS-APM 474
 |||:| ||||:||||:| : ||:|||| : ||| || |:|: || ||||| :
 Db 61 HLFNDGATVFFAVFMAVWATVFLEFWKRRRAVIAYDWDLIDWEEEEEEIRPQFEAKYSKK 120

Qy 475 TAPNPITGEDEPYFPERSRARRMLAGSVVIVVMVAVVVMCLVSIILYRAIMAIIVVSRGN 534
 |||:|: ||| : |:|: : | |: ||: : |:|:| :
 Db 121 ERMNPISGKPEPYQAFTDKCSRLIVSASGIFFMICVVIAAVFGIVIVRVVTV-----S 173

Qy 535 TLLA-AWA-----SRIASLTGSVV--NLVFILILSKIYVSLAHVLRWEMHRTQTKFEDA 586
 | | || |:|:| ||: | |:|:| :| :|: || | |:|:|:|:| :
 Db 174 TFAAFKWLIRNNSQVAT-TGTAVCINFCIIMLLNVLYEKVALLLTNLEQPRTESEWENS 232

Qy 587 FTLKVFIFQFVNFIYSSPVYIAFFKGRFVGYPGNYHTLFG-VRNEECAAGGLIELAQELL 645
 ||||:|:|||| || ||||| ||| |:| | | ||| ||||:| : :
 Db 233 FTLKMFLEFQFVNLNSSTFYIAFFLGRFTGHPGAYLRLINRWRLECHPSGCLIDLQMGMG 292

Qy 646 VIMVGKQVINNMQEVLPKLGWQKFLRLRSKKRKAGASAGASQGPWEDDYELVPCE--G 703
 :||| || || |: | : : || : | : : | | || ||| | |
 Db 293 IIMVLKQTWNNFMELGYPLIQNWWTR---RKVRQEHGPERKISFPQWEKDYNLQPMNAYG 349

RESULT 15
 US-10-066-543-1424
 ; Sequence 1424, Application US/10066543
 ; Publication No. US20030087818A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Jiang, Yuqiu
 ; APPLICANT: Pyle, Ruth A.
 ; APPLICANT: Xu, Jiangchun
 ; APPLICANT: Indirias, Carol Yoseph
 ; APPLICANT: Lodes, Michael J.
 ; APPLICANT: Secrist, Heather
 ; APPLICANT: Carter, Darrick
 ; APPLICANT: Fanger, Gary R.
 ; APPLICANT: Smith, Carole L.
 ; APPLICANT: Durham, Margarita
 ; APPLICANT: Stolk, John A.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 ; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
 ; FILE REFERENCE: 210121.563
 ; CURRENT APPLICATION NUMBER: US/10/066,543
 ; CURRENT FILING DATE: 2002-01-31
 ; NUMBER OF SEQ ID NOS: 3417
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 1424
 ; LENGTH: 782
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-066-543-1424

Query Match 21.0%; Score 1037.5; DB 4; Length 782;
 Best Local Similarity 33.3%; Pred. No. 7.4e-91;
 Matches 284; Conservative 137; Mismatches 282; Indels 149; Gaps 28;

```

Qy      111 DFVLVWEEDLKLDRQQDSAARDRTDMHRTWRETFLDNLRAAGLCV----DQQDVQDGNNT 166
          |::|: : |::| || : :|: || | : |::| |
Db      35 DYVLVAQRHTQRDPQ---AR-----QQQFLEELRRKGFIKVIDRQKQVFFG--- 79

Qy      167 VHYALLSASWAVLCYYAEDLRKLPLQELPNQASNWSAGLLAWLGIPNVLLLEVVPDPVPE 226
          : | : | | | | | | : : |
Db      80 -----IRADNSVFGLY----RTLLLEPEGPAPHAELAA-----PTTIPV 114

Qy      227 YYSFRFRVKNLPRFLGSDNQDTFFTSTKRHQILFEILAKTPYGHEKKNLLGIHQLLAEGV 286
          | | | : | : : | | | | | | | : | | | : | |
Db      115 TTSLRIRI---VNFVVMNN-----KTSAGET-----FEDLMKDG 146

Qy      287 LSAAPFLHDGPFKTPPEGPQAPRLNQQLFQHWARWGKWNKYQPLDHVRRYFGEKVALY 346
          | | | | | | | | | | | : | | | : | | | | | | |
Db      147 FEARFPLHKG-----EG----RLKKT-----WARWRHMFREQPVDEIRNYFGEKVALY 190

Qy      347 FAWLGFYTGWLLPAAVVGTFLVFLVGCFLVFSDIPTQELCGSKDSFEMCPL-CLDCPFWLL 405
          | | | | | | | | | | | | | : : | : | | | | : |
Db      191 FVWLGWYTYMLVPAALTGLLVFLSGFSLFEASQISKEICEAHDIL-MCPLGDHSRRYQRL 249

Qy      406 SSACALAQAGRLFDHGGTVFFSLFMALWAVLLEYYWKRKSATLAYRWDCSDYEDTEERPR 465
          | | | : | | | : | | | | | : | | | : | | : : |
Db      250 SETCTFAKLTHLFDNDGTVVFAIFMALWATVFLEIWKQRARVVLHWDLYVWDEEQEEMA 309

Qy      466 PQFAASAPMTAPNPITGEDEPYFPERSRARRMLAGSVVIVVMVAVVMCLV-----SIIL 520
          | : | : | : | : | : | : | : | : | : | : | : |
Db      310 LQLI-----NCPDYKLRPYQHSYLRSTVILV--LTLLMICLMIGMAHVLV 353

Qy      521 YRAIMAIVVSRSGNTLLAAWASRIASLTGSVNVLVFILILSKIYVSLAHVLTWRWEMHRTQ 580
          | | : : | | | : : | : | : | : | : | : | : | : |
Db      354 YRVLASALFSSSAVPFLEEQTAVVVTGALVHYVTIVIMTKINRRVALKLCDFEMPRTF 413

Qy      581 TKFEDAFTLKVFIFQVNFYSSPVYIAFFKGRFVGYPGNYHTLFGV-RNEECAAGGCLIE 639
          : : | | : | | : | : | | | | | : | : | : | | : |
Db      414 SERESRFTIRFFTLQFFTHFSSLIYIAFILGRINGHPGKSTRLAGLWKLECHASGCMMD 473

Qy      640 LAQELLVIMVGKQVINNMQEVLPKLGWQKFRLRSKKRKAGASAGASQGP---WEDD 695
          | : : | | : | : | : | : | : | : | : | : | : |
Db      474 LFMQMAIIMGLKQTLNCVEYLV-----WVTHKCRS--LRASESGHLPRDPELRDWRN 526

Qy      696 YELVPCE--GLFDEYLEMVLQFGFVTIFVAACPLAPLFAALLNNWVEIRLDARKFVCEYRR 753
  
```

```

      | | |      | | | : | : | : | : | | | | | | | | | : | | | | | | | | |
Db      527 YLLNPVNTFSLFDEFMEMMIQYGFTTIFVAAFPLAPLLALFSNLVEIRLDAIKMVWLQRR 586

Qy      754 PVAERAQDIGIWFHILAGLTHLAVISNAFLAFSSDFLPRAYRW-----TRAHD 803
      | : | : | | | : | : | | | : | : | : | : | : | : | : | : | : |
Db      587 LVPRKAKDIGTWLQVLETIGVLAVIANGMVIAFTSEFIPRVVYKYRYSPLKEGNSTVDC 646

Qy      804 LRGFLNFTLA-----RAPSSFAAAHNRT-CRYRAFRD-DDGHYSQTYWNLLAIRLAFV 854
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      647 LKGYVNHSLSVFHTKDFQDPDGIEGSENVTLCRYRDYRNPDPYNFSEQFWLLAIRLAFV 706

Qy      855 IVFEHVVFVSVGRLLDLLVPDIPESVEIKV-KREYYLAKQALAENEVLFGTNGTKDEQPKG 913
      | : | : | | : : | | | : | : | : | : | : | : | : | : | : | : |
Db      707 ILFEHVALCIKLIAAWFVPDIPQSVKNKVLEVKYQRLREKMWHRQRLGGVGAGSRPP-- 764

Qy      914 SELSSHWTPFTV 925
      : : | | : :
Db      765 --MPAHPTPASI 774

```

Search completed: October 27, 2006, 20:33:54
Job time : 195 secs

SCORE 1.3 BuildDate: 12/06/2005

SCORE Search Results Details for Application 10552515 and Search Result us-10-552-515-1.rapbn.

[Score Home Page](#)
[Retrieve Application List](#)
[SCORE System Overview](#)
[SCORE FAQ](#)
[Comments / Suggestions](#)

This page gives you Search Results detail for the Application 10552515 and Search Result us-10-552-515-1.rapbn.

[start](#)

[Go Back to previous page](#)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2006, 20:30:53 ; Search time 47 Seconds
(without alignments)
1646.261 Million cell updates/sec

Title: US-10-552-515-1
Perfect score: 4950
Sequence: 1 MRMAATAWAGLQGPPPLTLC.....SELSSHWPFTVPKASQLQQ 933

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 316380 seqs, 82930642 residues

Total number of hits satisfying chosen parameters: 316380

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA_New:*
1: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
2: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
3: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
4: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
5: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
6: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
7: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
8: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4950	100.0	933	6	US-10-552-515-1 Sequence 1, Appli
2	905	18.3	642	7	US-11-293-697-4483 Sequence 4483, Ap
3	684.5	13.8	483	7	US-11-293-697-3990 Sequence 3990, Ap
4	594.5	12.0	660	7	US-11-293-697-3644 Sequence 3644, Ap
5	411.5	8.3	393	7	US-11-197-712-457 Sequence 457, App
6	408.5	8.3	876	6	US-10-449-902-42265 Sequence 42265, A
7	353	7.1	466	6	US-10-449-902-50939 Sequence 50939, A
8	283.5	5.7	516	6	US-10-953-349-5638 Sequence 5638, Ap
9	283.5	5.7	543	6	US-10-953-349-5637 Sequence 5637, Ap

10	283.5	5.7	558	6	US-10-953-349-5636	Sequence 5636, Ap
11	105	2.1	1089	6	US-10-196-749-266	Sequence 266, App
12	104	2.1	697	6	US-10-449-902-49799	Sequence 49799, A
13	102	2.1	442	7	US-11-056-355B-8075	Sequence 8075, Ap
14	101	2.0	529	6	US-10-449-902-43353	Sequence 43353, A
15	101	2.0	532	6	US-10-449-902-52268	Sequence 52268, A
16	100	2.0	453	6	US-10-805-394-6877	Sequence 6877, Ap
17	100	2.0	472	6	US-10-449-902-53989	Sequence 53989, A
18	100	2.0	3010	7	US-11-140-487A-770	Sequence 770, App
19	96.5	1.9	477	6	US-10-449-902-51479	Sequence 51479, A
20	96.5	1.9	477	6	US-10-449-902-56002	Sequence 56002, A
21	96	1.9	258	6	US-10-953-349-27870	Sequence 27870, A
22	96	1.9	258	7	US-11-056-355B-65384	Sequence 65384, A
23	95.5	1.9	460	7	US-11-056-355B-70833	Sequence 70833, A
24	95.5	1.9	517	7	US-11-056-355B-70832	Sequence 70832, A
25	95.5	1.9	572	7	US-11-056-355B-70831	Sequence 70831, A
26	95	1.9	530	7	US-11-296-657-10	Sequence 10, Appl
27	95	1.9	735	7	US-11-366-965-799	Sequence 799, App
28	95	1.9	3010	7	US-11-140-487A-769	Sequence 769, App
29	93.5	1.9	466	7	US-11-434-137-2794	Sequence 2794, Ap
30	93.5	1.9	466	7	US-11-434-184-2794	Sequence 2794, Ap
31	93.5	1.9	466	7	US-11-434-199-2794	Sequence 2794, Ap
32	93.5	1.9	466	7	US-11-434-203-2794	Sequence 2794, Ap
33	93.5	1.9	790	6	US-10-449-902-52717	Sequence 52717, A
34	93	1.9	928	7	US-11-056-355B-107694	Sequence 107694, A
35	93	1.9	928	7	US-11-056-355B-118933	Sequence 118933, A
36	93	1.9	945	7	US-11-056-355B-107693	Sequence 107693, A
37	93	1.9	945	7	US-11-056-355B-118932	Sequence 118932, A
38	93	1.9	971	7	US-11-056-355B-107692	Sequence 107692, A
39	93	1.9	971	7	US-11-056-355B-118931	Sequence 118931, A
40	92.5	1.9	424	7	US-11-056-355B-17569	Sequence 17569, A
41	92.5	1.9	510	6	US-10-449-902-31185	Sequence 31185, A
42	92.5	1.9	613	7	US-11-056-355B-17568	Sequence 17568, A
43	92.5	1.9	635	7	US-11-056-355B-17567	Sequence 17567, A
44	92.5	1.9	703	7	US-11-056-355B-71193	Sequence 71193, A
45	92.5	1.9	715	7	US-11-056-355B-71192	Sequence 71192, A

ALIGNMENTS

RESULT 1

US-10-552-515-1

; Sequence 1, Application US/10552515

; Publication No. US20060194204A1

; GENERAL INFORMATION:

; APPLICANT: The Government of the United States of America as

; APPLICANT: represented by the Secretary of the Department of Health and

; APPLICANT: Human Services

; APPLICANT: Bera, Tapan K.

; APPLICANT: Pastan, Ira H.

; APPLICANT: Lee, Byungkook

; TITLE OF INVENTION: GENE EXPRESSED IN PROSTATE CANCER AND METHODS OF USE

; FILE REFERENCE: 4239-68223-02

; CURRENT APPLICATION NUMBER: US/10/552,515

; CURRENT FILING DATE: 2005-10-06

; PRIOR APPLICATION NUMBER: PCT/US2004/10588

; PRIOR FILING DATE: 2004-04-05

; PRIOR APPLICATION NUMBER: 60/461,399

; PRIOR FILING DATE: 2003-04-08

; NUMBER OF SEQ ID NOS: 12

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 1

; LENGTH: 933

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Splice Variant-Novel Gene Expressed in Prostate

US-10-552-515-1

Query Match 100.0%; Score 4950; DB 6; Length 933;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 933; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRMAATAWAGLQGPPLPTLCPAVRTGLYCRDQAHAEWAMTSETSSGSHCARSMLRRRA 60
|||||

```

Db      1 MRMAATAWAGLQGPPLPTLCPAVRTGLYCRDQAHAEWAMTSETSSGSHCARSRMLRRRA 60
Qy      61 QEEDSTVLIDVSPPEAEKRGSYGSTAHASEPGGQQAACRAGSPAKPRIADFLVWEEEDL 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 QEEDSTVLIDVSPPEAEKRGSYGSTAHASEPGGQQAACRAGSPAKPRIADFLVWEEEDL 120
Qy      121 KLDROQDSAADRTDMHRTWRETFLDNLRAAGLCVDQDQDGNNTTVHYALLSASWAVLC 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 KLDROQDSAADRTDMHRTWRETFLDNLRAAGLCVDQDQDGNNTTVHYALLSASWAVLC 180
Qy      181 YYAEDLRLKLPLQELPNQASNWSAGLLAWLGINVLLLEVVPDVPPEYSCRFRVKNLPRF 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 YYAEDLRLKLPLQELPNQASNWSAGLLAWLGINVLLLEVVPDVPPEYSCRFRVKNLPRF 240
Qy      241 LGSNDQDTFFTSTKRHQILFEILAKTPYGHEKKNLLGIHQLLAEGVLSAAFPLHDGPFKT 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 LGSNDQDTFFTSTKRHQILFEILAKTPYGHEKKNLLGIHQLLAEGVLSAAFPLHDGPFKT 300
Qy      301 PPEGQPAPRLNQOVLFQHWARWGKWNKYQPLDHVRRYFGEKVALYFAWLGFYTGWLLPA 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 PPEGQPAPRLNQOVLFQHWARWGKWNKYQPLDHVRRYFGEKVALYFAWLGFYTGWLLPA 360
Qy      361 AVVGTLVFLVGCFLVFSDIPTQELCGSKDSFEMCPLCLDCPFWLLSSACALAQAGRLFDH 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 AVVGTLVFLVGCFLVFSDIPTQELCGSKDSFEMCPLCLDCPFWLLSSACALAQAGRLFDH 420
Qy      421 GGTVFVFSLFMALWAVLLLEYWKRKSATLAYRWDCSDYEDTEERPRPQFAASAPMTAPNPI 480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 GGTVFVFSLFMALWAVLLLEYWKRKSATLAYRWDCSDYEDTEERPRPQFAASAPMTAPNPI 480
Qy      481 TGEDEPYFPERSRARRMLAGSVVIVVMVAVVMCLVSIILYRAIMAIVVSRSGNTLLAAW 540
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      481 TGEDEPYFPERSRARRMLAGSVVIVVMVAVVMCLVSIILYRAIMAIVVSRSGNTLLAAW 540
Qy      541 ASRIASLTGSVNLVLFILILSKIYVSLAHVLTREWEMHRTQTKFEDAFTLKVFIFQFVNFY 600
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      541 ASRIASLTGSVNLVLFILILSKIYVSLAHVLTREWEMHRTQTKFEDAFTLKVFIFQFVNFY 600
Qy      601 SSPVYIAFFKGRFVGYPGNYHTLFGVRNEECAAGGCLIELAQELLVIMVGKQVINNMQEV 660
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      601 SSPVYIAFFKGRFVGYPGNYHTLFGVRNEECAAGGCLIELAQELLVIMVGKQVINNMQEV 660
Qy      661 LIPKLKGWQKFRRLRSKKRKAGASAGASQGPWEDDYELVPCEGLFDEYLEMVLQFGFVTI 720
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      661 LIPKLKGWQKFRRLRSKKRKAGASAGASQGPWEDDYELVPCEGLFDEYLEMVLQFGFVTI 720
Qy      721 FVAACPLAPLFAALLNNWEIRLDARKFVCEYRRPVAERAQDIGIWFHILAGLTHLAVISN 780
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      721 FVAACPLAPLFAALLNNWEIRLDARKFVCEYRRPVAERAQDIGIWFHILAGLTHLAVISN 780
Qy      781 AFLFASDDFLPRAYYRWTRAHDLRGFLNFTLARAPSSFAAAHNRTCryRAFRDDGHYS 840
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      781 AFLFASDDFLPRAYYRWTRAHDLRGFLNFTLARAPSSFAAAHNRTCryRAFRDDGHYS 840
Qy      841 QTYWNLAILRALFVIVFEHVVSFVGRLLDLLVPDIPESVEIKVKREYYLAKQALAENEVL 900
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      841 QTYWNLAILRALFVIVFEHVVSFVGRLLDLLVPDIPESVEIKVKREYYLAKQALAENEVL 900
Qy      901 FGTNGTKDEQPKGSELSSHWPFTVPKASQLQQ 933
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      901 FGTNGTKDEQPKGSELSSHWPFTVPKASQLQQ 933

```

```

RESULT 2
US-11-293-697-4483
; Sequence 4483, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458

```

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4483
; LENGTH: 642
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-293-697-4483

Query Match 18.3%; Score 905; DB 7; Length 642;
Best Local Similarity 35.4%; Pred. No. 2.9e-73;
Matches 222; Conservative 106; Mismatches 219; Indels 80; Gaps 17;

```
Qy      26. GLYCRDQHAERWAMT--SETSSGSHCARSRMLRRRAQEEDSTVLIDVSPPEAEKRGSYG 83
      ||| || : : : || : | | | | | : |
Db      24. GLYFRDGRRKVDYILVYHHKRPSG-----NRTLVRVQHS DTP-----SGA 64

Qy      84. STAHASEPGGQAAACRAGSPAKPRIADFVLVWEEDLKLDROQDSAARDRTDMHRTWRET 143
      : | : || : || | : | : | : | : |
Db      65. RSVKQDHP L PGKASLDAGSGEPP-----MDYHEDD-----KRFRREE 102

Qy     144. FLDNLRAAGLCVDQDQDVGNTTVH---YALLSASWAVLCYYAEDLRLKPLQELPNQAS 200
      : || || : : : | : | : : | | | | | : | : | : | : |
Db     103. YEGNLL EAGLELE----RDEDTKIHG VGFVKIHAPWNVLCREAEFLKLMPTKKMYH--I 156

Qy     201. NWSAGLLAWLGIPNVLLEVVPDPPEYYSCR-----FRV NKLPRFLGSDNQDTFF 250
      | : || | | : || : : : | : | | | | | | | | | : | : |
Db     157. NETRGLLK--KINSVLQKITDPIQPKVAEHRPQTMKRLSYFPSREKQHLFDLSD-KDSFF 213

Qy     251. TSTKRHQILFEILAKTPYGHEKNLLGIHQLLAEGVLSAAFLHDGPFKTPPEGPQAPRL 310
      | | | : || | : | : : || | | | : || : || | | :
Db     214. DSKTRSTIVYEILKRTTCTKAKYS-MGITSLANGVYAAAYPLHDGDY-----NGENVEF 267

Qy     311. NQRQVLFQHWARWGKWKYQPLDHVRRYFGEKV ALYFAWLGFYTGWLLPAAVVGT LVLV 370
      | : : | : | | : | : | : | : | : | : | : | : | : | : |
Db     268. NDRKLLYEAWARYGVFYKYQPIDLVRKYFGEKIGLYFAWLGVYTQMLIPASIVGIIVFLY 327

Qy     371. GCFLVFSDIPTQELCGSKDSFEMCPLC-LDCPFWLLSSACALAQAGRLFDHGGTVFFSLF 429
      || : : || : | : : | || | | : | : || | | : | : || | : |
Db     328. GCATMDENIPSMEMCDQRHNITMCPLCDKTC SYWKMSACATARASHLFDNPATVFFSVF 387

Qy     430. MALWAVLLLEYWKRSATLAYRWDCSDYEDTEERPRPQFAA-----SAPMTAPNPITGED 484
      || | | : : || | | | | : : | : | : | : | : | : | : |
Db     388. MALWAATFMEHWKRKQMRNLNWRDLTGFEEDHPRAEYEARVLEKSLKKESRNKET--D 445

Qy     485. EPYFPERSRARRMLAGSVVIVVMVAVVMCLVSIILYRAIMAIVVSRSGNTLLAAWASRI 544
      : | | | | | : | : | : | : | : | : | : | : | : | : |
Db     446. KVKLTWRDRFPAYLTNLVSIIFMIAVTFAIVLGVIIRISMAAALAMNSSPSVRSNIRVT 505

Qy     545. ASLTGSVNVNLVFI LILSKIYVSLAHVLTRWEMHRTQTKFEDAFTLKVFIFQFVNFYSSPV 604
      : | : || | : : | : : | : | : | : | : | : | : | : |
Db     506. VTATAVIINLVV ILLDEVYGC IARWLTKIEVPKTEKSFEERLIFKAFLLKFVNSYTPIF 565

Qy     605. YIAFFKGRFVGYPGNYHTLF-GVRNEE 630
      | : || | | | | | | : | : | : |
Db     566. YVAFKGRFVGRPGDYVYIFRSFRMEE 592
```

RESULT 3
US-11-293-697-3990
; Sequence 3990, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3990
; LENGTH: 483
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-293-697-3990

Query Match 13.8%; Score 684.5; DB 7; Length 483;
 Best Local Similarity 35.1%; Pred. No. 1.9e-53;
 Matches 171; Conservative 95; Mismatches 168; Indels 53; Gaps 14;

```

Qy      479 PITGEDE-----PYFPERSARRMLAGSVVIVVMVAVVMCLV-----SIILYRAIM 525
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db      2  PAVSEEEMALQLINCPDYKLRPYQHSYLRSTVILV--LTLLMICLMIGMAHVLVVRVLA 59

Qy      526 AIVVSRSGNTLLAAWASRIASLTGSVVNLVFILILSKIYVSLAHVLTRWEMHRTQTKFED 585
      :  :  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db      60  SALFSSSAVPFLEEQVTTAVVVTGALVHYVTIVIMTKINRRVALKLCDFEMPRTFSERES 119

Qy      586 AFTLKVFIFQFVNFYSSPVYIAFFKGRFVGYPGNYHTLFGV-RNEECAAGGCLIELAQEL 644
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db      120 RFTIRFFTLQFFTHFSSLIYIAFILGRINGHPGKSTRLAGLWKLECHASGCMMDLQVQM 179

Qy      645 LVIMVGKQVINNMQEVLPKLGWQKFRRLRSKKRKAGASAGASQGP----WEDDYELVP 700
      :  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db      180 AIIMGLKQTLNSCVEYLVP-----WVTHKCRS--LRASEGHLPRDPELRDWRNRYLLNP 232

Qy      701 CE--GLFDEYLEMVLQFGFVTIFVAACPLAPLALLNNWVEIRLDARKFVCEYRRPVAER 758
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db      233 VNTFSLFDEFMEMMIQYGFTTIFVAAFLAPLALLFNSNLVEIRLDAIKMVWLQRRLLVPRK 292

Qy      759 AQDIGIWFHILAGLTHLAVISNAFLAFSSDFLPRAYYRW-----TRAHDLRGFL 808
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db      293 AKDIGTWLQVLETIGVLAVIANGMVIATSEFIPRVVYKYRYSPLKEGNSTVDCLKGYV 352

Qy      809 NFTLA-----RAPSSFAAAHNRT-CRYRAFRD-DDGHYSQTYWNLLAIRLAFVIVFEH 859
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db      353 NHSLSVFTKDFQDPDGIEGSENVTLCRYRDRNPPDYNFSEQFWLLAIRLAFVILFEH 412

Qy      860 VVFSVGRLLDLLVPDIPESVEIKV-KREYYLAKQALAENEVLFGTNGTKDEQPKGSELSS 918
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db      413 VALCIKLIAAWFVPDIPQSVKNKVLEVQYQLREKMWHRQRLGGVGAGSRPP----MPA 468

Qy      919 HWTPTTV 925
      |  |  |  |
Db      469 HPTPASI 475
  
```

RESULT 4

US-11-293-697-3644

```

; Sequence 3644, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3644
; LENGTH: 660
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-293-697-3644
  
```

Query Match 12.0%; Score 594.5; DB 7; Length 660;
 Best Local Similarity 24.1%; Pred. No. 4.1e-45;
 Matches 171; Conservative 122; Mismatches 231; Indels 185; Gaps 20;

```

Qy      228 YSCRFRVKNLPRFLG-SDNQDTFFTSTKRHQILFEILAKTPYGHEKKNLLG----- 277
      :  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db      100 FTYRTRQN----FKGFDNNDFFLTMAECQFII-----KHELENLRKDEKMIPGY 146

Qy      278 -----IHQLLAEGVLSAAFPPLHDGPFKTPPEGPQAPRLNQRQVLFQHW-ARWGK 325
      :  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db      147 PQAKLYPGKSLRLRLTSGIVIQVFLHDS-----EALKKLEDTWYTRFAL 192

Qy      326 WNKYQPLDHHVRRYFGEKVALYFAWLGFYTGWLLPAAVVGTLVFLVGCFLVFSDIPTQELC 385
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db      193 --KYQPIDSIRGYFGETIALYFGFLEYFTFALIPMAVIG----- 229
  
```



```

Qy      386 GSKDSFEMCPLCLDCPFWLLSSACALAQAGRLFDHGGTVFFSLFMALWAVLLLEYWKRKS 445
          |::|          |: | | : | :| :| | |
Db      230 -----LPYYLFVWE-----DYDKYVIFASFNLIWSTVILELWKRGC 265

Qy      446 ATLAYRWDCSDYEDTEERPRPQFAASAPMTAPNPITGEDEPYFPERSRARRMLAGSVVIV 505
          | : | | | : | | | | : | | | : | | : | : |
Db      266 ANMTYRWGTLLMKRKFEEPRPGFHG---VLGINSITGKEEPLYPSYKRLRIYLVSLPFV 322

Qy      506 VMVAVVMCLVSIILYRAIMAIVVSRSGNTLLAAWASRIASLTGSVVNLVFILILSKIYV 565
          : : : | : | : : : | | : | : : | : : |
Db      323 CLCLYFSLYVMMIYFDMEVWALGLHENS---SEWTS-VLLYVPSIIYAIVIEIMNRLYR 378

Qy      566 SLAHLVLRWEMHRTQTKFEDAFTLKVFIFQFVNFFYSSPVYIAFFKGRFVGYPGNYHTLFG 625
          | | | | | : : : | | : | : | : | | |
Db      379 YAAEFLTSWENHRLSAYQNHLILKVLVFNFLNCFASLFYIAFV----- 422

Qy      626 VRNEECAAGGCLIELAQELLVIMVGKQVINNMQEVLIPLKLGW---QKFRRLRSKKRKAGA 683
          : : : | | | : : | : | : | : | : | : |
Db      423 LKDMKL-----LRQSLATLLITSQILNQIMESFLP----YWLQRKHGVRVVRKVQAL 470

Qy      684 SAGASQGPWED---DYELVPCEGFLFDEYLEMVLQFGFVTIFVAACPLAPLFAALLNNWVEI 740
          | : | : | : | | : | | : | | : | | : |
Db      471 KADIDATLYEQVILEKEMGTLYGTDDYLELFLQFGYVSLFSCVYPLAAFAVLNNFTEV 530

Qy      741 RLDARKFVCEYRRPVAERAQDIGIWFHILAGLTHLAVISNAFLAFSSDFLPRAYRWTR 800
          | | | : | | : | : | : | : : | | : | : :
Db      531 NSDALKMCRVFKRPFSEPSANIGVWQLAFETMSVISVVTNCALIGMSPQV--NAVFPESK 588

Qy      801 AHDLRGFLNFTLARAPSSFAAAHNRTCRYRAFRDDGHYSQTYWNLLAIRLAFVIVFEHV 860
          | | | : | : | : | : | : | : | : | : |
Db      589 A-DL-----ILIVVAVEHA 601

Qy      861 VFSVGRLLDLLVPDIPESVEIKVKREYYLAKQALAENEVLFGTNGTKDE 909
          : : : : | : | : | : | : | : : | : |
Db      602 LLALKFILAFAPDKPRHIQMKLARLEFESLEALKQQQMKLVTENLKEE 650

```

RESULT 5

US-11-197-712-457

; Sequence 457, Application US/11197712

; Publication No. US20060130160A1

; GENERAL INFORMATION:

; APPLICANT: Dumas Milne Edwards, Jean Baptiste

; APPLICANT: Bougueleret, Lydie

; APPLICANT: Jobert, Severin

; TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS

; FILE REFERENCE: 78.US4.CIP

; CURRENT APPLICATION NUMBER: US/11/197,712

; CURRENT FILING DATE: 2005-08-04

; PRIOR APPLICATION NUMBER: US/09/876,997

; PRIOR FILING DATE: 2001-06-08

; PRIOR APPLICATION NUMBER: US 09/731,872

; PRIOR FILING DATE: 2000-12-07

; PRIOR APPLICATION NUMBER: US 60/187,470

; PRIOR FILING DATE: 2000-03-06

; PRIOR APPLICATION NUMBER: US 60/169,629

; PRIOR FILING DATE: 1999-12-08

; NUMBER OF SEQ ID NOS: 482

; SOFTWARE: Patent.pm

; SEQ ID NO 457

; LENGTH: 393

; TYPE: PRT

; ORGANISM: Homo sapiens

US-11-197-712-457

Query Match 8.3%; Score 411.5; DB 7; Length 393;

Best Local Similarity 23.8%; Pred. No. 7.4e-29;

Matches 111; Conservative 95; Mismatches 172; Indels 89; Gaps 11;

```

Qy      448 LAYRWDCSDYEDTEERPRPQFAASAPMTAPNPITGEDEPYFPERSRARRMLAGSVVIVM 507
          : | | | : | | | | : | | | : | | : | : |
Db      1 MTYRWGTLLMKRKFEEPRPGFHG---VLGINSITGKEEPLYPSYKRLRIYLVSLPFVCL 57

Qy      508 VAVVMCLVSIILYRAIMAIVVSRSGNTLLAAWASRIASLTGSVVNLVFILILSKIYVSL 567
          : : : | : | : : : | | : | : : | : : |

```

```

Db      58 CLYFSLYVMMIYFDMVWALGLHENS---SEWTS-VLLYVPSIIYAIVIEIMNRLYRYA 113
Qy      568 AHVLRWEMHRTQTKFEDAFTLKVFIFQFVNFYSSPVYIAFFKGRFVGYPGNYHTLFGVR 627
      | | | | | : : : | | : | : : | | | : :
Db      114 AEFLTSWENHRLESAYQNHILKVLVFNFLNCFASLFYIAFV-----LK 157
Qy      628 NEECAAGGCLIELAQELLVIMVGKQVINNMQEVLI PKLKGW--QKFRLSKRRKAGASA 685
      : : | | : : : | : : | : | : | : : |
Db      158 DMKL-----LRQSLATLLITSQILNQIMESFLP----YWLQRKHGVRVVRKRVQALKA 205
Qy      686 GASQGPWED---DYELVPCGLFDEYLEMVLQFGFVTIFVAACPLAPLFA LLNNWVEIRL 742
      : | : | : | | | | : | | | | : | :
Db      206 DIDATLYEQVILEKEMGTYLGTDDYLEFLQFGYVSLFSCVYPLAAFAVLNNFTEVNS 265
Qy      743 DARKFVCEYRRPAERAQDIGIWFHILAGLTHLAVISNAFLAFSSDFLPRAYRWTRAH 802
      | | | : | | : : | : : : | : | : : |
Db      266 DALMKCRVFKRPFSEPSANIGWQLAFETMSVISVVTNCALIGMSPOV--NAVFPESKA- 322
Qy      803 DLRGFLNFTLARAPSSFAAAHNRTCryRAFRRDDGHYSQTYWNLLAIRLAFVIVFEHVVF 862
      | | : | : | :
Db      323 DL-----ILIVVAVEHALL 336
Qy      863 SVGRLLDLLVPDIPESVEIKVKREYYLAKQALAENEVLFGTNGTKDE 909
      : : | | : : | : : : | : : : | : |
Db      337 ALKFILAFAPDKPRHIQMKLARLEFESLEALKQQQMKLV TENLKEE 383

```

RESULT 6

US-10-449-902-42265

; Sequence 42265, Application US/10449902

; Publication No. US20060123505A1

; GENERAL INFORMATION:

; APPLICANT: National Institute of Agrobiological Sciences.

; APPLICANT: Bio-oriented Technology Research Advancement Institution.

; APPLICANT: The Institute of Physical and Chemical Research.

; APPLICANT: Foundation for Advancement of International Science.

; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF

; FILE REFERENCE: MOA-A0205Y1-US

; CURRENT APPLICATION NUMBER: US/10/449,902

; CURRENT FILING DATE: 2003-05-29

; PRIOR APPLICATION NUMBER: JP 2002-203269

; PRIOR FILING DATE: 2002-05-30

; PRIOR APPLICATION NUMBER: JP 2002-383870

; PRIOR FILING DATE: 2002-12-11

; NUMBER OF SEQ ID NOS: 56791

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 42265

; LENGTH: 876

; TYPE: PRT

; ORGANISM: Oryza sativa

US-10-449-902-42265

Query Match 8.3%; Score 408.5; DB 6; Length 876;

Best Local Similarity 22.4%; Pred. No. 4.2e-28;

Matches 198; Conservative 110; Mismatches 285; Indels 291; Gaps 39;

```

Qy      73 PP-----EAEK--RGSYGSTAHASEPGGQQAACRAGSPAKPRIADFLVWEEDLK--L 122
      | | : : | | : : | : | : | : | : |
Db      28 PPIQSALHSAQKAADAVGSARDATSSSSRTTEYPVPQSHPDQIADYVLVFQHI AKKYL 87
Qy      123 DRQQDSAARDRTDMHRTWRETFLDNLRAAGLCVDQQDVQDGNTTVHYALLSASWAVLCYY 182
      | : | : : : : : | | : : | : | :
Db      88 RSSTKVPAAERSKIAAEY-DALVQRIRDTGLHVT SREGSKG-----SGQILLFV 135
Qy      183 AEDRLRLKPLQELPNQ--ASNWSAGLLAWLGI PNVLLEVVPDPVPEYYSRFRVNKL--P 238
      | : | : | : | : | : | : | : | :
Db      136 KADSQL---LHQLARQEALSDYLHGVL S-----VQPPPP-----RSSSLADA 174
Qy      239 RFLGSDNQDTFFTSTKRHQILFEILA-KTPYGHEKNLLGIHQLLAEGVLSAA---FP-- 292
      | | : : : : : | | | : : | |
Db      175 SFQKQQPSSLHLTPASRLRLVDSLLTLP SLASHAVKNAAGSSQVPSGAGRLGLKEFPHL 234
Qy      293 -----LHDGPFKTPPEGFPQAPRLNQRQVLFQHWARWGKWNKYQP-----LDHVR 336
      : | | | : : | : : | : |
Db      235 VDMSAIHD-----PAYNSA-----WMK--RWSHTSPAKLFSGIGLADLDSIR 274

```

Qy	337	RYFGEKVALYFAWLGFYTGWLLPAAVVGTLVFLVGCFLVFSIDIPTQELCGSKDSFEMCPL	396
Db	275	EHFGEDVALYFGFLNFYFQALAPAALLG-----	302
Qy	397	CLDCPFWLLSSACALAQAGRLFDHGGTVFFSLFMALWAVLLLEYWKRKSATLAYRWDCSD	456
Db	303	---LAFWLL-----GRPF---SPVYSLGLVSVSCLFVELWRMKERKLAVRWGTLG	346
Qy	457	YEDTEERPRPQFAASAPMTAPNPITGEDEPYFPERSRARRMLAGSVVI-----VVMVAVV	511
Db	347	VSQVDRR-RHDFVPRT--TRIDPATEEPEEVFEWWRRELRVILSLPIVAFFASVLAATMT	403
Qy	512	VMCLVSII---LYR-----AIMAIVVSRSGNTLLAAWASRIASLTGSVVNLVF	556
Db	404	LMFVVEIFVTQLYHGPLKQAVPFIPTALLVVAVPQ---IIAAWQ-----	444
Qy	557	ILILSKIYVSLAHVLTWRWEMHRTQTKFEDAFTLKVFIFQFVNFY-----SSPVYI-----	606
Db	445	-----ATAVAITKWNHHSAKSYDYSLTLRKFAMQAITAYGALTLSAYIYIPFGES	495
Qy	607	-----AFFKGR-----FVGYPGNYHTLFGVRNEECAAGGLIELAQ	642
Db	496	IMETMVQRGFFRDSIQHAIRTGNISPKGIDFHINPDRMHT-----	535
Qy	643	ELLVIMVGKQVINNMQEVLI PKLKGWQKFR-----LRSKKRKAGASA----	685
Db	536	QLFAVSVTSQFVNAFTELALPVLMRKFAEWRGERAAQKNADGSSQRQQRQDSASSAPS	595
Qy	686	-----GASQGPWED-----DYELVPC EGLFDEYLEMVLQFGFVTIFVAAC	725
Db	596	SGDVTPLEPGAADGGKEESERRFVSRVRKELQLPPYD-LFGDYAEMATQFGYITLWSVW	654
Qy	726	PLAPL FALLNNWVEIRLDARKEFVCEYRRPVAERAQDIGIWFHILAGLTHLAVISNA---F	782
Db	655	PLSPVMG FVNFFELRSDAAKISVNNRRPVPVRAETIGPWLEAFGFIAWL GALNNAALVY	714
Qy	783	LLAFSSDFLPRAYYRW-----TRAHDLRGLNFTLA-RAPSSFA	820
Db	715	LFOOSEHAHLEGHSRYETTMRTLHPSRA--NLTLADDAQSRFS	756

```

Db      47 LPA-----FFFFVISWAVFLQFWKRKNSAVLARWGINYSFSEYKTMG 89
Qy      462 ERPRPQFAASAPMTAPNPITGEDEPYFFPERSRARRMLAGSVVI-----VVMVAVVVMCL 515
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      90 NELDFLSFSMADDNVQQRKFGAPK----EKSIVQRNEWFGVLLRIRNNAIIVLAIICLQL 145
Qy      516 VSIILYRAIMAIIVSRSGNTLLAAWASRIASLTGSSVNLVFILILSKIYVSLAHVLRWE 575
      : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      146 PFELAYAHLYAITKTEALRYVLT-----VYLAAIQYTRIGGKVSVTLIKYE 193
Qy      576 MHRTQTKFEDAFTLKVFIQFVNIFYSSPVYIAFFKGRFVGYPGNYHTLFGVRNEECAAGG 635
      : : : | : | | | | | | | | | | | | | | | | | | | | | | |
Db      194 NNQGEQSSADSLVYKVF-----GLYFMQSYIGLF-----YHASLH-----RN----- 230
Qy      636 CLIELAQELLVIMVGKQVINNMQEVLPKLGWQWQKFR-LRSKKRKAGASAGAS---QGP 691
      : : | | : : | : | : | | | : | : | : | : | : | : | |
Db      231 -IMALRQVLIKRLIVSQVLENLIENSIPYLNYSYKKYRAVHKKKHEKESPAGKSVRLSTR 289
Qy      692 WEDDYELVPC-----EGLFDEYLEMVLQFGFVTIFVAACPLAPLFAALLNNWVEIR 741
      | : | | | : | | | : | | | : | | | : | | | : | | | : | | |
Db      290 VEKEY-LKPSYASIGEELEDGLFDDFLELTQFGMIMMFACAFPSIFCFAALNNVTEIR 348
Qy      742 LDARKFVCEYRRPVAERAQDIGIWFHILAGLTHLAVISNAFLFASSDFLPRAYYRWTRA 801
      | | | : : | | | | | | : | : | : | : | : | : | |
Db      349 ADALKLLVMLKRPAPRDAATIGAWLNIFQFLVMAICTNCLL----- 391
Qy      802 HDLRGFLNFTLARAPSSFAAAHNRTCRRYRAFRDDDDGHYSQTYWNLLAIRLAFVIVFEHV 861
      | | | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      392 -----VCLY-----DEEGKWK-----IEPLAAILIMEHAL 417
Qy      862 FSVGRLDLLVDPDIPESVEIKVKREYYLAKQALAE 897
      | : | : | : | : | : | : | : | : | : | : | : | : |
Db      418 FLIKFGFSHFVPE--EPAWVKANRGYVAQ---AQN 448

```

RESULT 8

US-10-953-349-5638

; Sequence 5638, Application US/10953349

; Publication No. US20060107345A1

; GENERAL INFORMATION:

; APPLICANT: ALEXANDROV, Nickolai et al.

; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES

; TITLE OF INVENTION: ENCODED THERBY

; FILE REFERENCE: 2750-1579PUS2

; CURRENT APPLICATION NUMBER: US/10/953,349

; CURRENT FILING DATE: 2004-09-30

; NUMBER OF SEQ ID NOS: 40252

; SOFTWARE: PatentIn version 3.3

; SEQ ID NO 5638

; LENGTH: 516

; TYPE: PRT

; ORGANISM: Arabidopsis thaliana

US-10-953-349-5638

Query Match 5.7%; Score 283.5; DB 6; Length 516;

Best Local Similarity 23.4%; Pred. No. 4.2e-17;

Matches 127; Conservative 78; Mismatches 167; Indels 171; Gaps 22;

```

Qy      422 GTVFFSLFMALWAVLLEWKRKSATL----AYRWDCSDYEDTEERPRPQFAASAPMTAP 477
      | | : | | | | | : | | | | | | | | : | : | : | : |
Db      108 GTI-----LWAALFLQFWKRKNAALLASQGYRFLGMEWSSL-----PFPKELIKNLG 154
Qy      478 NPITGEDEPY-----FPERSRARRMLAGSVVIVVMVAVVVMCLVSIILYRAIMAIIVSRS 532
      | | | | | | | | : | : | : | : | : | : | : | : | : | : |
Db      155 NERAKEKEAYQRYEWFAYRKFRN-----DVLVIMSIICLQLPFELAYAHIFEIITS- 207
Qy      533 GNTLLAAWASRIASLTGSSVNLVFILILSKIYVSLAHVLRWEMHRTQTKFEDAFTLKVF 592
      : : | : | : | : | : | : | : | : | : | : | : | : |
Db      208 -----IIKYVLTAIYLLIIQYLTR-----LGGKVSVKLI 236
Qy      593 ---IFQFVNIFYSSPVYIAFFKGRFVGYPGNYHTLFGVRNEECAAGGCLIELAQELLVIMV 649
      | : | : : : : | | | | | | : | | | : : | : | : : |
Db      237 NREINESVEYRANSLIYKTYIGIF-----YHVLLH-RN-----FMTLRQVLIQRLI 281
Qy      650 GKQVINNMQEVLPKLGWQWQKFRRLRSKKR-KAGASAGASQ--GPWEDDY----- 696
      | | : : : | | : | : | : | : | : | : | : | : | : |

```

```

Db      282 ISQVFWTLMGSLPYLKYSYRKYRARTKKKMGDSSTGKIQIASRVEKEYFKPTYSASIG 341
Qy      697 -ELVPCEGLFDEYLEMVLQFGFVTIFVAACPLAPLFAALLNNWEIRLDARKFVCEYRRPV 755
      || :|||: ||: ||| : :| | ||| | :| :||| :| | : |||:
Db      342 VELE--DGLFDDSLALQFGMIMFACFPLAFALAAVSNVMEIRTNALKLLVTLRRPL 399
Qy      756 AERAQDIGIWFHILAGLTHLAVISNAFLAFSSDFLPRAYRWTRAHDLRGFLNFTLARA 815
      | ||| :| | :|| :|||
Db      400 PRAAATIGAWLNIWQFLVVMISICTNSALL----- 428
Qy      816 PSSFAAAHNRTCryAFRDDDGHYSQTYWNLAIRLAFVIVFEHVSVGRLLDLLVPDI 875
      | | | :| : : : | | :||| : : | |||:
Db      429 -----VCLY----DQEGKWK-----IEPLAAIIMEHVLLLLKFGLSRLVPEE 468
Qy      876 PESVEI-KVK-----REYLLAKQALAENEVLFGTNGTKDEQPKGSELSSHWPFTVPKAS 929
      | | :|| :| | ||| : : | : | | :
Db      469 PAWVRASRVKNVTQAQDMY-CKQLL-----RSISGEFNSLTKPEQE 508
Qy      930 QLQ 932
      | |
Db      509 QQQ 511

```

RESULT 9

```

US-10-953-349-5637
; Sequence 5637, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 5637
; LENGTH: 543
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-5637

```

```

Query Match          5.7%; Score 283.5; DB 6; Length 543;
Best Local Similarity 23.4%; Pred. No. 4.6e-17;
Matches 127; Conservative 78; Mismatches 167; Indels 171; Gaps 22;

```

```

Qy      422 GTVFFSLFMALWAVLLLEYWKRKSATL----AYRWDCSDYEDTEERPRPQFAASAPMTAP 477
      ||: ||| | | :|||: | | ||: :|
Db      135 GTI-----LWAALFLQFWKRKNAALLASQGYRFLGMEWSSL-----PFPKELIKNLG 181
Qy      478 NPITGEDEPY-----FPERSRARRMLAGSVVIVVMVAVVMCLVSIILYRAIMAVVSRS 532
      | | | | | | | | :|||: : | : | | | :|
Db      182 NERAKEKEAYQRYEFAYRKFRN-----DVLVIMSIIICLQLPFELAYAHIFEIITS- 234
Qy      533 GNTLLAASRIASLTGVSNNLVFILILSKIYVSLAHVLTREWEMHRTQTKFEDAFTLKVF 592
      : :| : ||: : ||| :||:
Db      235 -----IIKYVLTAIYLLIIQYLTR-----LGGKVSVKLI 263
Qy      593 ---IFQFVNFISSPVYIAFFKGRFVGYPGNYHTLFGVRNEECAAGGCLIELAQELLVIMV 649
      | :| : : : : | | | | | : | | : :
Db      264 NREINESVEYRANSLIYKTYIGIF-----YHVLLH-RN-----FMTLRQVLIQRLI 308
Qy      650 GKQVINNMQEVLPKLGWQKFRRLRSKKR-KAGASAGASQ--GPWEDDY----- 696
      || : : : | || :||: | ||| : :| | | | :|
Db      309 ISQVFWTLMGSLPYLKYSYRKYRARTKKKMGDSSTGKIQIASRVEKEYFKPTYSASIG 368
Qy      697 -ELVPCEGLFDEYLEMVLQFGFVTIFVAACPLAPLFAALLNNWEIRLDARKFVCEYRRPV 755
      || :|||: ||: ||| : :| | ||| | :| :||| :| | : |||:
Db      369 VELE--DGLFDDSLALQFGMIMFACFPLAFALAAVSNVMEIRTNALKLLVTLRRPL 426
Qy      756 AERAQDIGIWFHILAGLTHLAVISNAFLAFSSDFLPRAYRWTRAHDLRGFLNFTLARA 815
      | ||| :| | :|| :|||
Db      427 PRAAATIGAWLNIWQFLVVMISICTNSALL----- 455
Qy      816 PSSFAAAHNRTCryAFRDDDGHYSQTYWNLAIRLAFVIVFEHVSVGRLLDLLVPDI 875
      | | | :| : : : | | :||| : : | |||:

```

```

Db      456 -----VCLY----DQEGKWK-----IEPGLAAILIMEHVLLLLKFGLSRLVPEE 495
Qy      876 PESVEI-KVK-----REYYLAKQALAENEVLFGTNGTKDEQPKGSELSSHWPFTVPKAS 929
      | | :|| :| | | | :| : | | :
Db      496 PAWVRASRVKNVTQAQDMY-CKQLL-----RSISGEFNSLTKEPEE 535
Qy      930 QLQ 932
      | |
Db      536 QQQ 538

```

RESULT 10

```

US-10-953-349-5636
; Sequence 5636, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 5636
; LENGTH: 558
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-5636

```

```

Query Match          5.7%; Score 283.5; DB 6; Length 558;
Best Local Similarity 23.4%; Pred. No. 4.7e-17;
Matches 127; Conservative 78; Mismatches 167; Indels 171; Gaps 22;

```

```

Qy      422 GTVFFSLFMALWAVLLLEYWKRKSATL----AYRWDCSDYEDTEERPRPQFAASAPMTAP 477
      ||: ||| | |::|||:| | ||: :: |
Db      150 GTI-----LWAALFLQFWKRKNAALLASQGYRFLGMEWSSL-----PFPKELIKNLG 196
Qy      478 NPITGEDEPY-----FPERSRARRMLAGSVVIVVMVAVVMCLVSIILYRAIMAIVVSRS 532
      | | | | | | | | | | :|:|:|:| :| :| | | |
Db      197 NERAKEKEAYQRYEWFAYRKRFNRN-----DVLVIMSIIICLQLPFELAYAHIFEIITS- 249
Qy      533 GNTLLAAWASRIASLTGSVVNLVFIILSKIYVSLAHVLTWRWEMHRTQTKFEDAFTLKVF 592
      : :|: ||: : ||| :|:|:
Db      250 -----IIKYVLTAIYLLIIQYLTR-----LGKKVSVKLI 278
Qy      593 ---IFQFVNFYSSPVYIAFFKGRFVGYPGNYHTLFGVRNEECAAGGCLIELAQELLVIMV 649
      | :| :|:| :| :| | | | | | :| | | :| :|
Db      279 NREINESVEYRANSLIYKTYIGIF-----YHVLH-RN-----FMTLRQVLIQRLI 323
Qy      650 GKQVINNMQEVLIPLKLGWWQKFRRLRSKKR-KAGASAGASQ--GPWEDDY----- 696
      || : : :| || :|:| | |:: :|:| | | | :|
Db      324 ISQVFWTLMGSLPYLKYSYRKYRARTKKKMGDSSTGKIQIASRVEKEYFKPTYSASIG 383
Qy      697 -ELVPCEGLFDEYLEMVLQFGFVTIFVAACPLAPLFAALLNNWVEIRLDARKFVCEYRRPV 755
      || :|||: ||: ||| : :| | || | :| :||| :| | : |||:
Db      384 VELE--DGLFDDSLALQFGMIMFACAPLAFALAAVSNVMEIRTNALKLLVTLRRPL 441
Qy      756 AERAQDIGIWFHILAGLTHLAVISNAFLAFSSDFLPRAYYRWTRAHDLRGFLNFTLARA 815
      | | | :| | :|:|:| |
Db      442 PRAAATIGAWLNIWQFLVMSICTNSALL----- 470
Qy      816 PSSFAAAHNRTCRYRAFRDDDGHSYQTYWNLAIRLAFVIVFEHVVSFVGRLLDLLVPDI 875
      | | | :| : : | | :|:| :| :| | | :
Db      471 -----VCLY----DQEGKWK-----IEPGLAAILIMEHVLLLLKFGLSRLVPEE 510
Qy      876 PESVEI-KVK-----REYYLAKQALAENEVLFGTNGTKDEQPKGSELSSHWPFTVPKAS 929
      | | :|| :| | | | :| : | | :
Db      511 PAWVRASRVKNVTQAQDMY-CKQLL-----RSISGEFNSLTKEPEE 550
Qy      930 QLQ 932
      | |
Db      551 QQQ 553

```

RESULT 11

US-10-196-749-266
; Sequence 266, Application US/10196749
; Publication No. US20060094864A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C340
; CURRENT APPLICATION NUMBER: US/10/196,749
; CURRENT FILING DATE: 2002-07-16
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 266
; LENGTH: 1089
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-196-749-266

Query Match 2.1%; Score 105; DB 6; Length 1089;
Best Local Similarity 18.2%; Pred. No. 1.8;
Matches 108; Conservative 79; Mismatches 174; Indels 234; Gaps 27;

```

Qy      320 WARWGKWNKYQPLDHVRRYFGEKVALYFAWLGFYTGWLLPAAVVGTLVFLVGCFLVF--- 376
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      527 WKAWAGWGSKRPLATLFPPIGPVLLLLLFR LAVFFSDSFVVAEARATPFLGSGFILLLVV 586

Qy      377 -----SDIPTQELCGSKDS-----FEMC----P 395
      : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      587 QLHWEGQLLPPKLLTMPRLGTSATTNPPRHNGAYALRLGIGLLCTRLAGLFHRCPEETP 646

Qy      396 LCLDCPFWLLSSACALAQAGRLFDHGGTVFFSLFMALWAVLLLEYWKRKSATLAYRWDCS 455
      : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      647 VCHSSP-WLSPLASMVGGRAKNLWYGACV--AALVALLAAVRL--WLRRYGNL----- 694

Qy      456 DYEDTEERPRPQFAASAPMTAPNP-----ITGEDEPYFPERSRARRMLAG-SVVIVVMV 508
      : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      695 ---KSPEPPMLFVRWGLPLMALGTAAYWALASGADEA--PPRLRV--LVSGASMVLPRAV 747

Qy      509 AVVVMCLVSIILYRAIMAIVVSRSG-----NTLL----- 537
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      748 AGLAASGLALLLWKPVTVLVKAGAGAPRTRTVLTPFSGPPTSQADLDYVVPQIYRHMQUEE 807

Qy      538 -----AAW-----ASRIASLT-----GSVVNLVFILILSKIY 564
      | | : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      808 FRGRLERTKSQGPLTVAAYQLGSVSYAAMVTALTLLAFPLLLLHAERISLVFLLFLQSF 867

```



```
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 43353
; LENGTH: 529
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-43353
```

```
Query Match      2.0%; Score 101; DB 6; Length 529;
Best Local Similarity 20.6%; Pred. No. 1.5;
Matches 71; Conservative 53; Mismatches 137; Indels 84; Gaps 16;
```

```
Qy      463 RPRP---QFAASAPMTAP-----NPITGEDE---PYFPERSRARRMLAGS 501
      ||| :|:|:| | :|: || | : : | :
Db      50 RPRPPWCRFSASSPPPPDDDDYELDDTTGNCPLCSVDEVSSQYFEANYKPKNDLLKA 109

Qy      502 VVIVVMVAV-----VVMCLVSIILYRAIM---AIVVSRSGNTLLAA---W 540
      : | : : | | : | | : : : || | | |
Db      110 LTIIATALAGAAAINHSWVAEHQDIAMVLVFALGYAGIIFEESLAFNKSGVGLLMAVCLW 169

Qy      541 ASR-----IASLTGSVVNLVFIILSKIYVSLAHVLRWEMHRTQTKFEDAF 587
      | : : | | :| | : : | : : : :
Db      170 VIRSIGAPSTDVAVQELSHTTAEVSEIVFFLLGAMTIVEIVDAHQGFKLVTDNISTRNPR 229

Qy      588 TLKVFIQFQVNFY-----SSPVYIAFFKGRFVGYPGNYHTLFG-VRNEECAAGGC- 636
      || : : | | : : | : | : | | | | |
Db      230 TL-LWVIGFVTFFLSSILDNLTSTIVMVSLL--RKLVPPEYRKLGLGAVVVISANAGGAW 286

Qy      637 --LIELAQELLVIMVGKQVINNMQEVLIPLKLGWQKFRRLSKRKAGASAGASQGPWED 694
      : : : | | : | | : | : | : | |
Db      287 TPIGDVTTMLWIHQITTLNTMQGLFLPSVVS LAVPLALMSLTSEANGSSQKSSSLLSS 346

Qy      695 DYELVPCEGLFDEYLEMVLQFG-----FVTIFVAACPLAPL FALL 734
      : : | | : : | | : | | | | : :
Db      347 E-QMAP-RG-----QLVFAVGLGALVFVPVKALTGLPPFMGMM 383
```

RESULT 15

```
US-10-449-902-52268
; Sequence 52268, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 52268
; LENGTH: 532
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-52268
```

```
Query Match      2.0%; Score 101; DB 6; Length 532;
Best Local Similarity 20.6%; Pred. No. 1.5;
Matches 71; Conservative 53; Mismatches 137; Indels 84; Gaps 16;
```

```

Qy      463 RPRP---QFAASAPMTAP-----NPITGEDE---PYFPERSRARRMLAGS 501
      ||| :|:|:| | :|: || | : : | :
Db      50 RPRPPWCRFSASSPPPPDDDDYELDDTTGNCDPLCSVDEVSSQYFEANYKPKNDLLKA 109

Qy      502 VVIVVMVAV-----VVMCLVSIILYRAIM---AIVVSRSGNTLLAA---W 540
      : | : : | | : | | : : :| | | |
Db      110 LTIIATAGAAAINHSWVAEHQDIAMVLVFALGYAGIIFEESLAFNKSGVGLLMAVCLW 169

Qy      541 ASR-----IASLTGSVVNLVFILILSKIYVSLAHVLRWEMHRTQTKFEDAF 587
      | : : | | :| | : : | : : :
Db      170 VIRSIGAPSTDVAVQELSHTTAEVSEIVFFLLGAMTIVEIVDAHQGFKLVTDNISTRNPR 229

Qy      588 TLKVFIQFVNFY-----SSPVYIAFFKGRFVGYPGNYHTLFG-VRNEECAAGGC- 636
      || : : | | : : | : : | | | | |
Db      230 TL-LWVIGFVTFFLSSILDNLTSTIVMVSL--RKLVPPEYRKLLGAVVVISANAGGAW 286

Qy      637 --LIELAQELLVIMVGKQVINNMQEVLI PKLKGWWQKFR LRSKKRKAGASAGASQGPWED 694
      : : : | | : | | : | : | | : | :
Db      287 TPIGDVTTMLWIHGQITTLNMQGLFLPSVVSLAVPLALMSLTSEANGSSQKSSSLSS 346

Qy      695 DYELVPCEGLFDEYLEMVLQFG-----FVTIFVAACPLAPL FALL 734
      : : | | : | | | : | | | : :
Db      347 E-QMAP-RG-----QLVFAVGLGALVFPVFKALTGLPPFMM 383

```

Search completed: October 27, 2006, 20:34:47
Job time : 53 secs

SCORE 1.3 BuildDate: 12/06/2005

This page gives you Search Results detail for the Application 10552515 and Search Result us-10-55
start

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2006, 20:23:48 ; Search time 48 Seconds
(without alignments)
1870.213 Million cell updates/sec

Title: US-10-552-515-1
Perfect score: 4950
Sequence: 1 MRMAATAWAGLQGPPLPTLC.....SELSSHWTPTVTPKASQLQQ 933

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	734	14.8	1049	2	T22762	hypothetical prote
2	288.5	5.8	572	2	F96755	hypothetical prote
3	181.5	3.7	946	2	S48255	probable membrane
4	117	2.4	548	2	I48693	natural resistance
5	115.5	2.3	3010	1	GNWVTC	genome polypeptide
6	110.5	2.2	680	2	T35404	probable squalene-
7	110.5	2.2	873	2	S46584	probable membrane
8	110	2.2	792	2	T00487	probable potassium
9	108	2.2	3010	1	A45573	genome polypeptide
10	106.5	2.2	519	2	T11129	cytochrome-c oxida
11	106	2.1	438	2	B86088	probable citrate p
12	106	2.1	438	2	E91240	probable membrane
13	105	2.1	621	2	JC1346	dopamine beta-mono
14	104	2.1	646	2	H82555	c-type cytochrome
15	103.5	2.1	478	2	JQ2034	RNA-directed RNA p
16	102	2.1	302	2	C83993	hypothetical prote
17	101.5	2.1	395	2	D81040	cytochrome c-type
18	101.5	2.1	395	2	B81986	probable membrane
19	101	2.0	466	2	A95355	probable inner-mem
20	100.5	2.0	585	2	S74673	pleD protein - Syn


```

      : | | : | : | : | | | | | : : | | : : |
Db      272 TFLTLWPVIFGLVVFYIGFIDSISSAPLDWNHCKVNVFIGQTENVACGMRNGVTLFFS 331
Qy      393 MCPLCLDCPFWLLSSACALAQAGRLFDHGGTVFFSLFMALWAVLLEWKRKSATLAYRW 452
      | : | | : | | | | : : | | : : | : | : |
Db      332 MVTQ-----WFMSS-----FDTKMNAFFAVFMSIWGSVFVQIWKRNNSVLSYQW 375
Qy      453 DCSYEDTEERPRPQFAASAPMTAPNPITGEDEPYFPERSRARRMLAGSVVIVVMVAVV 512
      : | : | | : | | : | | | | | : | : | | : : | | |
Db      376 NSDDFHAIEP-DRPEFRGS--KVKEDPITGEDIWISPALARYIKMLASFVFSFMSLVV 432
Qy      513 MCLVSIILYRAIMAIVSRSGNTLLAAWASRIASLTGSVNVNLVFIILSKIYVSLAHVLT 572
      : | : : | : | : | | | : | : | : : | | : | |
Db      433 ISLMLVTLTKIWMVYNFQCTKEYTFHCWLS--AAFLPSILNTLSAMGLGAIYSNLVSRN 490
Qy      573 RWEMHRTQTQKFEDATLKVFIQFVNFYSSPVYIAFFKGRFVGYPGNYHTLFGVRNE--- 629
      | | | | : : : | : | | | | : | : | : | | : |
Db      491 SWENHRTESEHNNSLIVKIFAFQMVNTYTSIFYVAFIRPESHGLQPN--GLFGLGTEFKD 548
Qy      630 ECAAGGCLIELAQELLVIMVGKQVINNMQEVLP---KLKGWWQKFRRLRSKKRKAGASAG 686
      | | | | : | | : | : : | | | | | | | | |
Db      549 TCLDDTCSSLLALQLLTHLIKFPFKFDVLPYFVKL-----FRLRMYTSRTEARVE 602
Qy      687 ASQGPWEDDYELVPCEGLFDEYLEMVLQFGFVTIFVAACPLAPLALLNNWVEIRLDARK 746
      | | : | : | : | | | | | : | : | : | : |
Db      603 I-----EDDDQ-----ANVLMFASLFPLAPLALLIIGFVDMRIDAHR 639
Qy      747 FVCEYRRPVAERAQDIGIWFHILAGLTHLAVISNAFLAFSSDFLPRAYRWTRAHDLRG 806
      : | : | | | | | | : | : | : | : | : |
Db      640 LIWFNRKPIPIITNGIGIWLPIITFLQYCAVFTNAFIVAFSTSGFC-----684
Qy      807 FLNFTLARAPSSFAAAHNRTCARYAFRDDDGHYSQTYWNLLAIRLAFVIVFEHVVSFVGR 866
      | : | | | | | | | | | | | | | | : | :
Db      685 -----STFLA-----DGAYC-TVQN---RLIIVIVFQNLVFGKY 715
Qy      867 LLDLLVPDIPESVEIKVKREYY-----LAKQA-LAENEVLFGT 903
      | | : | | | : : : | : | : | : | : | :
Db      716 LLSSVIPSIPASIKLALRKKRYVVAHIVEKGDVPHRTRIKRTRIAKLAWIASNQMKIS 775
Qy      904 NGTKDEQPKGSELSSHW 921
      | | : | | : |
Db      776 NRKKEKSNK-----KHFT 788

```

RESULT 2

F96755

hypothetical protein F3N23.22 [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004

C;Accession: F96755

R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, J.; Altaf, H.; Ara Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.J.; Koo, H.L.; Krem

A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, L.J.; Tambunga, G.;

A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: F96755

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-572

A;Cross-references: UNIPROT:Q9SSM5; UNIPARC:UPI00000A63BD; GB:AE005173; NID:g5903091; PIDN:AAD55649.1; GSPD

C;Genetics:

A;Gene: F3N23.22

A;Map position: 1

```

Query Match          5.8%; Score 288.5; DB 2; Length 572;
Best Local Similarity 20.5%; Pred. No. 3.4e-16;
Matches 169; Conservative 97; Mismatches 236; Indels 323; Gaps 29;

```

```

Qy      142 ETFLDNLRAAGLCVDQDQDVQDGNNTTVHYALLSASWAVLCYAE DLRLKLPQLPQASN 201
      | : | | | : | | | | | | | | | | | |
Db      32 EVLVTELRRKKGMVVDR-----VVGLAHEFLKVAAPSEILGNAAAE 71

```

```

Qy      202 WSAGLLAWLGIPNVLLVVPDPPEYYSCFRVKNLPRFLGSDNQDTFTSTKRHQILFE 261
      | | | | | | | | | | | | | | | |
Db      72 LHIRKPTRLGI-----DLPFEMQGEAFIRQPDGLLFS-----WFERFRCYQHLYI 117

```

```

Qy      262 ILAKTPYGHEKKNLLG-----IHQLLAEGVLSAAPFLHDGPFKTPPEGPQAPR 309
      : : ||: | : | : ||: ||||
Db      118 GIVNSG-GHDVTLKLDGREFCWTAGESLLRRLESEGVKQMFPLHDE----- 163

Qy      310 LNQRQVLQFQHWA-RWGKWN-KYQPLDHVRRYFGEKVALYFAWLGFYTGWLLPAAVVGTIV 367
      :|: | |:|| | || ||:| : ||| |
Db      164 -LKRKELLQNWALNW--WNCTNQPIDQIYSYFGAK----- 195

Qy      368 FLVGCFLVFSDIPTQELCGSKDSFEMCPLCLDCPFWLLSSACALAQAGRLFDHGGTVFFS 427
      | : |
Db      196 -----ELIKNLGN----- 203

Qy      428 LFMALWAVLLLEYWKRKSATLAYRWDSCSYEDTEERPRPQFAASAPMTAPNPITGEDEPY 487
      | | | | |
Db      204 -----ERAKEKEAYQRYEW----- 217

Qy      488 FPERSRARRMLAGSVVIVVMVAVVVMCLVSIILYRAIMAIVVSRSNTLLAAWASRIASL 547
      | | | | :|:|:|: | : | | | :| |
Db      218 FAYRKRFRN-----DVLVIMSIICLQLPFELAYAHIFEIITSDIKYVLT----- 263

Qy      548 TGSVVNLVFIILISKIYVSLAHVLTREWEMHRTQTKEFADFTLKVFIFQFVNFYSSPVYIA 607
      : | : | :|: | :|: | :| : || | |
Db      264 ----IYLLIIQYLTRLGGKVSVKLINREINESVEYRANSLIYKVF-----GLYFMQTYIG 314

Qy      608 FFKGRFVGYPGNYHTLFGVRNEECAAGGLIELAQELLVIMVGKQVINNMQEVLIPLKLG 667
      | | | | | : | | | : | | : : | |
Db      315 IF-----YHVLLH-RN-----FMTLRQVLIQRLIISQVFWTILMDGSLPYLKY 355

Qy      668 WWQKFRRLRSKKR-KAGASAGASQ--GPWEDDY-----ELVPC EGLFDEYLEMVL 713
      :|:| |:|: | :| | | :| :|||: | :|
Db      356 SYRKYRARTKKKMGDSSTGKIQIASRVEKEYFKPTYSASIGVELE--DGLFDDSLAL 413

Qy      714 QFGFVTIFVAACPLAPLFAALLNNWVEIRLDARKFVCEYRRPVAERAQDIGIWFHILAGLT 773
      ||| : :| | || | :| :||: | | : |||: | | | :|
Db      414 QFGMIMFACAFPLAFALAAVSNVMEIRTNALKLVTLRRPLPRAAATIGAWLNIWQFLV 473

Qy      774 HLAVISNAFLAFSSDFLPRAYRWTRAHDLRGFLNFTLARAPSSFAAAHNRTCRYAFR 833
      : : :|: || | |
Db      474 VMSICTNSALL-----VCLY----- 488

Qy      834 DDDGHYSQTYWNLLAIRLAFVIVFEHVVSVGRLLDLLVPDIPESVEI-KVK-----REY 887
      | : | : | | :| :||: | : | ||: | | :|| : :
Db      489 DQEGKWK-----IEPGLAAILIMEHVLLLLKFGLSRLVPEPAWVRASRVKNVTQAQDM 542

Qy      888 YLAKQALAENEVLFGTNGTKDEQPKGSELSSHWTPTVPKASQLQ 932
      | | | | : | : | | : | |
Db      543 Y-CKQLL-----RSISGEFNSLTKEQEQQQ 567

```

RESULT 3

S48255

probable membrane protein YBR086c - yeast (*Saccharomyces cerevisiae*)

N;Alternate names: hypothetical protein YBR0809

C;Species: *Saccharomyces cerevisiae*

C;Date: 03-Aug-1995 #sequence_revision 11-Aug-1995 #text_change 09-Jul-2004

C;Accession: S48255; S45954; S44670

R;Mannhaupt, G.; Stucka, R.; Ehnle, S.; Vetter, I.; Feldmann, H.

Yeast 10, 1363-1381, 1994

A;Title: Analysis of a 70 kb region on the right arm of yeast chromosome II.

A;Reference number: S48255; MUID:95208357; PMID:7900426

A;Accession: S48255

A;Status: nucleic acid sequence not shown

A;Molecule type: DNA

A;Residues: 1-946

A;Cross-references: UNIPROT:P38250; UNIPARC:UPI0000036C25; EMBL:X78993; NID:g476045; PIDN:CAA55593.1; PID:g

R;Feldmann, H.; Mannhaupt, G.; Schwarzlose, C.; Vetter, I.

submitted to the Protein Sequence Database, August 1994

A;Reference number: S45927

A;Accession: S45954

A;Molecule type: DNA

A;Residues: 1-946

A;Cross-references: UNIPARC:UPI0000036C25; EMBL:Z35955; NID:g536351; PID:g536352; MIPS:YBR086c

C;Genetics:

A;Gene: SGD:IST2

A;Cross-references: SGD:S0000290

A;Map position: 2R
C;Superfamily: Saccharomyces cerevisiae probable membrane protein YBR086c
C;Keywords: transmembrane protein
F;131-147/Domain: transmembrane #status predicted
F;158-174/Domain: transmembrane #status predicted
F;207-243/Domain: transmembrane #status predicted
F;248-274/Domain: transmembrane #status predicted
F;302-324/Domain: transmembrane #status predicted
F;450-477/Domain: transmembrane #status predicted
F;506-532/Domain: transmembrane #status predicted
F;563-588/Domain: transmembrane #status predicted

Query Match 3.7%; Score 181.5; DB 2; Length 946;
Best Local Similarity 18.6%; Pred. No. 8.6e-07;
Matches 118; Conservative 99; Mismatches 243; Indels 173; Gaps 25;

```

Qy      342 KVALYFAWLGFYTGWLLPAAVGTLVLVVGCFLVFSDIPTQELCGSKDSFEMCPLCLDCP 401
          | :|||:| | ||: | : : |
Db      119 KQSLYFAFLQNYIKWLIPFSFFGLSIRFLSNF----- 150

Qy      402 FWLLSSACALAQAGRLFDHGGTVFFSLFMAWVLLLEYWKRK----SATLAYRWDCSD 456
          :: | :||| || : :| | | :|
Db      151 -----TYEFNST--YSLFAILWTLSTAFWLYKYEPFWSDRLSKYSSFST 193

Qy      457 YEDTEERPRPQFAASAPMTAPN----PITGEDEPYFPERSRARRMLAGSVVIVVMVAVV 512
          | :: : | ||: : | : :| | : : : :
Db      194 IEFLQDKQKAQKKASSVIMLKCCFIPVA-----LLFGA----ILLSFQL 234

Qy      513 MCLVSIILYRAIMAIIVSRSGNTLLAAWASRIASLTGSVVNLVFILILSKIYVS-LAHVL 571
          | | :| :| : : | | : : | :| || :
Db      235 YCFALEIFYKQIY-----NGPMI-----SILSFLPTILICTFTPVLTVIYNKYFVEPM 282

Qy      572 TRWEMHRTQTKFEDAFTLKVFIFQFVNFYSSPVYIAFFKGRFVGYPGNYHTLFGVRNEEC 631
          | :| | : : : | :| : : | :| :| : : :| :
Db      283 TKWENHSSVVNAKKSKEAKNFVIFLSSY-VPLLITL----FLYLPMGHLLTAEIRTKVF 337

Qy      632 AAGGCLIEL-----AQELLVIMVGKQVINNMQEVLIPLKLGWWQK----- 671
          | | | : : :| :| | :| | :| | :|
Db      338 NAFSILARLPTHDSDFIIDTKRYEDQFFYFIVINQLIQFSMENFVPSLVSIAQQKINGPN 397

Qy      672 ---FRLRSKKRKAGASAGASQGPWE---DDYELVPCGLFD---EYLEMVLQFGFVTIFVA 723
          : | : || | : : | :| | || : : :||| : :|
Db      398 PNFVKAESSEIGKAQLSS-SDMKIWSKVKSQYQTDWPGATFDLDANFKLLQLQGYLVMFST 456

Qy      724 ACPLAPLFAALLNNWVEIRLDARKFVC----EY-RRPVAERAQD-----IGIWFHILA 770
          |||| | : | : :| || | || | : : : :| :| :|
Db      457 IWPLAPFICLIVNLIVYQVDLRKAVLYSKPEYFPFIYDKPSSVSNTQKLTVGLWNSVLV 516

Qy      771 GLTHL-AVISNAFLAFSSDFLPRAYRWTRAHDLRGFLNFTLARAPSSFAAAHNRTCRY 829
          : | || : : | :| : : : :| :| |
Db      517 MFSILGCVITATLTMYQSCNIP-----GVGAHTSIHTNKAWY 554

Qy      830 RAFRDDDGHYSQTYWNLIAIRLAFVIVFEHVVSFVGRLLDLLVPDIPESVEIKVKREYYL 889
          | : : :| : : : ||| : : | : : :| : : :
Db      555 LA-----NPINHSWINI----VLYAVFIEHVSVAIFFLSSILKSSHDDVANGIVPKHV 605

Qy      890 AKQALAENEVL-----FGTNGTKD-EQPKGS 914
          | : || : :| :| :| |||
Db      606 NVQNPPKQEVFEKIPSEFNSNEKELVQRKGS 638

```

RESULT 4
I48693
natural resistance-associated macrophage protein 1 - mouse
N;Alternate names: transport system membrane protein Nramp
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C;Accession: I48693; A57071; A40739
R;Barton, C.H.; White, J.K.; Roach, T.I.; Blackwell, J.M.
J. Exp. Med. 179, 1683-1687, 1994
A;Title: NH2-terminal sequence of macrophage-expressed natural resistance-associated macrophage protein (Nr
A;Reference number: I48693; MUID:94216838; PMID:7513015
A;Accession: I48693
A;Status: preliminary; translated from GB/EMBL/DBBJ
A;Molecule type: mRNA
A;Residues: 1-548.

A;Cross-references: UNIPROT:P41251; UNIPARC:UPI000002770F; EMBL:X75355; NID:g505155; PIDN:CAA53102.1; PID:g
R;Govoni, G.; Vidal, S.; Cellier, M.; Lepage, P.; Malo, D.; Gros, P.
Genomics 27, 9-19, 1995
A;Title: Genomic structure, promoter sequence, and induction of expression of the mouse Nramp1 gene in macr
A;Reference number: A57071; MUID:95394476; PMID:7665187
A;Accession: A57071
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-548
A;Cross-references: UNIPARC:UPI000002770F
R;Vidal, S.M.; Malo, D.; Vogan, K.; Skamene, E.; Gros, P.
Cell 73, 469-485, 1993
A;Title: Natural resistance to infection with intracellular parasites: isolation of a candidate for Bcg.
A;Reference number: A40739; MUID:93258812; PMID:8490962
A;Accession: A40739
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 65-548
A;Cross-references: UNIPARC:UPI0000178BFA
A;Note: sequence extracted from NCBI backbone (NCBIN:131666, NCBIP:131667)
C;Superfamily: natural resistance-associated macrophage protein 1

Query Match 2.4%; Score 117; DB 2; Length 548;
Best Local Similarity 20.2%; Pred. No. 0.13;
Matches 136; Conservative 89; Mismatches 207; Indels 240; Gaps 34;

```

Qy      67 VLIDVSPPEAEKRGSYGSTAHASEPG-GQQAACRAGSPAKPRIADFLVWEEIDLKLDRO 125
      : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1 MISDKSPRL-SRPSYSGSI--SSLPGPAPQPAPCR-----ETYLSEKIP 41

Qy     126 QDSAARDRTDMHRTWRET-----FLD--NLRAAGLCVDQDDVDGNTTVHYALLSA 174
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db     42 IPSADQGTFSRLKLWAF TGPGFILMSIAFLDPGNI-----ESDLQAGAVAGFKLLWVL 93

Qy     175 SWA----VLC-----YYAEDLRLKL-----PLQELPNQ 198
      || : || : : : : : : : : : : : : : : : : : : : : :
Db     94 LWATVGLLQCRLAARLGVVTGKDLGEVCHLYYPKVPRILLWLTIELAIVGSDMQEVIGT 153

Qy     199 ASNW---SAGLLAWLGIPNVLLVVPDVPPEYYSCRFRVNKLPRFLGSDNQDTFFTSTKR 255
      | : : | | : | : | : | : | : : : : | : : : | | |
Db     154 AISFNLLSAGRIPLWG--GVLITIV-DTFFFLDNYGLRKLEAFFG----- 197

Qy     256 HQILFEILAKTPYGHEKKNLLGIHQLLAEGVLSAAFPLHDGPFKTPPEGPQAPRLNQRQV 315
      : | : | : | : | : | : | : | : | : | : | : | : | :
Db     198 --LLITIMALT-FGYE---YVVAHP--SQGALLKGLVLP TCGCGQPELLQAVGIVGAI 249

Qy     316 LFQHWARWGKWNKYQPLDHVRRYFGEKVALYFAWLGFYTGWLLPAAVVGTLVFLVGCFL- 374
      : : : : : : : : : : : : : : : : : : : : : : : : :
Db     250 MPHNIYLHSALVKSREVDRTRRVDVREANMYF-----LIEATIALSVSFIINLFVM 300

Qy     375 -VFSDIPTQELCGSKDSFEMCPLCLDCPFWLLSSACALAQAGRLFD-----HGG 422
      || : : : : : : : : : : : : : : : : : : : : : ||
Db     301 AVFGQAFYQQT--NEEAFNIC-----ANSSLQNYAKIFPRDNNTVSVDIYQGG 346

Qy     423 TVFFSLF---MALWAVLLEYWKRKSATLAY-----RWDCSDYEDTEERPRP 466
      : || : : | | : : | | : : | | : : | | : : | |
Db     347 VILGCLFGPAALYIWAVGLLAAGQSSTMTGTAGQFVMEGFLKLRW----- 392

Qy     467 QFAASAPMTAPNPITGEDEPYFPERSR-ARRMLAGSVVIV--VMVAV-----V 511
      : : : : : : : : : : : : : : : : : : : : : : :
Db     393 -----SRFARVLLTRSCAILPTVLVAVFRDLKDL SGLNDL 427

Qy     512 VMCLVSIIYRAIMAI VSRSGNTLLAAWAS-RIASLTGS-----VVNLVFILILSKI 563
      : | : : | : : | : : : : : : : : : : : : : : : :
Db     428 LNLVQSLLLPFAVLPILTFTSMPAVMQEFANGRMSKAITSCIMALVCAINLYFVI---SY 484

Qy     564 YVSLAHVLTRWEMHRTQTKFEDAFTLKVFIFQFVNFYSSPVYIAF----FKGRFVGYPG 618
      || | : : : : : : : : : : : : : : : : : : : :
Db     485 LPSLPH-----PAYFGLVALFA-IGYLGLTAYLAWTCCIAHGATFLTHSS 528

Qy     619 NYHTLFGVRNEE 630
      : | | : | |
Db     529 HKHFLYGLPNEE 540

```

RESULT 5

GNWVTC

genome polyprotein - hepatitis C virus

N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructural protein NS3); m

C;Species: hepatitis C virus

C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 09-Jul-2004

C;Accession: A38465

R;Takamizawa, A.; Mori, C.; Fuke, I.; Manabe, S.; Murakami, S.; Fujita, J.; Onishi, E.; Andoh, T.; Yoshida, J. Virol. 65, 1105-1113, 1991

A;Title: Structure and organization of the hepatitis C virus genome isolated from human carriers.

A;Reference number: A38465; MUID:91140698; PMID:1847440

A;Accession: A38465

A;Molecule type: genomic RNA

A;Residues: 1-3010

A;Cross-references: UNIPROT:P26663; UNIPARC:UPI0000131E1C; EMBL:M58335; NID:g329770; PIDN:AAA72945.1; PID:g

C;Superfamily: hepatitis C virus genome polyprotein

C;Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstructural protein; nucleoti

F;2-115/Product: capsid protein C #status predicted

F;116-191/Product: envelope protein M #status predicted

F;192-389/Product: major envelope protein E #status predicted

F;390-729/Product: nonstructural protein NS1 #status predicted

F;730-1006/Product: nonstructural protein NS2 #status predicted

F;1007-1615/Product: hepacivirin #status predicted

F;1230-1237/Region: nucleotide-binding motif A (P-loop)

F;1312-1317/Region: nucleotide-binding motif B

F;1316-1319/Region: DEXH motif

F;1616-1862/Product: nonstructural protein NS4a #status predicted

F;1863-2013/Product: nonstructural protein NS4b #status predicted

F;2014-3010/Product: nonstructural protein NS5 #status predicted

F;196,209,234,250,305,325,417,423,430,448,532,540,556,576,623,645,1213,1255,2041,2077,2240,2529,2788/Bindin

Query Match 2.3%; Score 115.5; DB 1; Length 3010;

Best Local Similarity 19.7%; Pred. No. 1.7;

Matches 163; Conservative 83; Mismatches 285; Indels 295; Gaps 40;

```
Qy      128 SAARDRTDMHRTWRETFLDNLRAAGLCVDQ-----QDVQDGNTTVHYALLSASWAVLCYY 182
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      314 SGHRMAWMMNWSPT-----TALVVSQLLRIPQAVVDMVAGAHWGVL----AGLAYY 362

Qy      183 AEDLRLKLPLQELPNQASNWSAGLLAWLGIPNVLLLEVVPDVPPEYYSCRFRVKNLPRFLG 242
      : | | | : | : | | | | | | | | | | | | | | | |
Db      363 -----SMAGNWKVLIVML-----LFAG 380

Qy      243 SDNQDTFFT-----STKRHQILF-----EILAKTPYGHEKKNLLGIHQLLAEGVLS 288
      | | | | | | | | | | | | | | | | | | | | | |
Db      381 VDG-DTHVTGGAQAKTTNRLVSMFASGPSQKIQLINTNGSWHINRTALNCNDSLQTGFLA 439

Qy      289 AAFPLHDGPFKTPPE-----GP-----QAPRLNQRQVLFQH----- 319
      | | | | | | | | | | | | | | | | | | | | | |
Db      440 ALFYTHSFNSGCPERMAQCRTIDKFDQGWGPITYAESSRSDQRPYCWHPYPPQCTIVPA 499

Qy      320 -----WARWGK-----WNKYQPLDHVRRYFGEKVALY 346
      | | | | | | | | | | | | | | | | | | | | | |
Db      500 SEVCGPVYCFTSPVVVGTTDRFGVPTWRGENETDVLLNNTRPPQ--GNWFG----- 551

Qy      347 FAWL---GFYTGWLLPAAVVG-----TLVFLVGCFLVFSDIPTQELCGSKDSFEMCPLCL 398
      | : | | | | : | | | | | | | | | | | | | | | |
Db      552 CTWNNSTGFTKTCGGPPCNIGVGNNLTLCPTDCFRKHPE-ATYTKCGSGP--WLTPRCM 608

Qy      399 -DCPFWLLSSACALAQAGRLFDHGGTVFFSLFMALWAVLLLEYWKRKSATLAYRWDCSDY 457
      | | : | | : | | | | | | | | | | | | | | | |
Db      609 VDYPY-----RLWHYPCTVNFITFKVRMYVGGVEH--RLNA--ACNWTRGER 651

Qy      458 EDTEERPRPQFAASAPMTAPNPITGEDEPYFPERSRARRMLAGSVVIVVMVAVVMCLVS 517
      | | : | | : | | | | | | | | | | | | | | | |
Db      652 CDLEDRLPELSPLLLSTTEWQVLPSCFTTLPALSTGLIHLHQINVDVQYLYGIGSAVVS 711

Qy      518 IIL---YRAIMAIVVSRSGNTLLAAWAS-RIASLTGSSVNLVLFILILSKIYVSLAHVLR 573
      : | : : : : | | | | | | | | | | | | | | | |
Db      712 FAIKWEYVLLFLLLA-DARVCACLWMLLIAQAEAALENLV---VLNSASVAGAH---- 763

Qy      574 WEMHRTQTKFEDAFTLKVFIFQFVNFYSSPVYIAFFKGRFVGYPGNYHTLFGV----- 626
      | | : | : | | | | | | | | | | | | | | | |
Db      764 -----GILSFLVFCAAWYI---KGRLV--PGATYALYGWVPLLLLL 800

Qy      627 -----RNEECAAGGCLIELAQELLVIMVGKQV--INNMQEVLIPKLKGWQKFR 673
      | : | | | | | | | | | | | | | | | | | | | |
```

```

Db      801 LALPPRAYAMDREMAASCGG-----AVFVGLVLLTSPYYKVFLARLIWWLQYFT 850
Qy      674 LRSKKRKAGASAGASQGPW-----EDDYELVPC---EGLFDEYLEMVLQFGFVTI 720
      |         | |         |         | : |         | : |         | : :
Db      851 TR-----AEADLHVWIPPLNARGGRDAIILLMCAVHPELIFDITKLLIAILGPLMV 901
Qy      721 FVAACPLAPLALLNNWVEIRLDARKFVCEYRRPVAERAQDIGIWFHILAGLT-----H 774
      |         | |         : |         | | |         : : | |         |
Db      902 LQAGITRVPYF-----VRAQGLIHACMLVRKVA-GGHYVQMAFMKLGALTGTYYINH 952
Qy      775 LAVISNAFLAFSSDFLPRAYRWTRAHDLRGFLNFTLARAPSSFA 820
      | : :         | | |         | : : | | :
Db      953 LTPLRD-----WPRA---GLRDLAVAVEPVVFS 977

```

RESULT 6

T35404

probable squalene-hopene cyclase - *Streptomyces coelicolor*

C;Species: *Streptomyces coelicolor*

C;Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004

C;Accession: T35404

R;Oliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A. submitted to the EMBL Data Library, March 1999

A;Reference number: Z21577

A;Accession: T35404

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-680

A;Cross-references: UNIPROT:Q9X7V9; UNIPARC:UPI00000DAF47; EMBL:AL049485; PIDN:CAB39697.1; GSPDB:GNO0070; S

A;Experimental source: strain A3(2)

C;Genetics:

A;Gene: SCOEDB:SC6A5.13

C;Superfamily: squalene-hopene cyclase

Query Match 2.2%; Score 110.5; DB 2; Length 680;
 Best Local Similarity 22.8%; Pred. No. 0.61;
 Matches 112; Conservative 52; Mismatches 173; Indels 155; Gaps 27;

```

Qy      40 MTSETSSGSHCARSRMLRRRAQEEDSTVLIDVSPPEAEKRGSYGSTAHASEPGGQAAAC 99
      ||: ||: || | | | | | | | | | | | | | | | | | | | | | | | |
Db      1 MTA-TTDGSTGASLRPLAASASDTDITI-----PAAAAGVPEAAA- 39
Qy     100 RAGSPAKPRIADFVLV-----WEEDLKLDRQQDSAARDRTDMHRTWRETFL-----DN 147
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      40 ----RATRATDFLLAKQDAEGWKKGD-----ETNVTMDAEDL---LLRQFLGIQDEET 87
Qy     148 LRAAGLCVDQDQDVGNTTVHY-ALLSASWAVLCYYAEDLRLKLPLQELPN--QASNW-- 202
      ||| | | : : || | | | | | | | | | | | | | | | | | | | |
Db      88 TRAAALFIRGEQREDGTWATFYGGPGELSTTIEAYVA--LRLAGDSPEAPHMARAAEWIR 145
Qy     203 SAGLLA-----WLGIPN-VLLEVVPDPVPE--YYSRFRVKNLPRFLGSDNQDTFFT 251
      | | : | | | | | | | | | | | | | | | | | | | | | | | |
Db     146 SRGGIASARVFTRIWLALFGWKKWDDLPELPPELIYF-----PTWVPLNIYD--FG 194
Qy     252 STKRHQI--LFEILAKTPYGHEKKNLLGIHQLLAEGVLSAAFP---LHDGPFKTPPEGPQ 306
      | | | | : | | | | | | | | | | | | | | | | | | | | |
Db     195 CWARQTIIVPLTIVSAKRP-----VRPAPFPLDELHTDPARPNPPRPL 236
Qy     307 AP-----RLNQ-----RQVLFQHWARW-----GKWNKYQPLDHVR 336
      ||         | : :         | :         |||         | | | |
Db     237 APVASWDGAFQRIDKALHAYRKVAPRRLRAAMNSAARWIIERQENDGCWGGIQP----- 291
Qy     337 RYFGEKVVALYFAWLGFYTGWLLPAAVVGTLVFLVGCFLVFSDIPTQELCGSKDSFEMCPL 396
      | : |         : | : | | : : : | | : :         | | |
Db     292 -----PAVYSVIALYLLGYDLEHPVMRAGLES�DRFAVWRE-----DGARMIEA 335
Qy     397 CLDCPFW-LLSSACALAQAGRLFDHGGTVFFSLFMALWAVLLLEYWKRKSATL-----AY 450
      | | | | : | | | | | | | | | | | | | | | | | | | | |
Db     336 C-QSPVWDTCLATIALADAGVPEDHPQLVKASDWMLGEQIVRPGDWSVKRPGLPPEGWAF 394
Qy     451 RWDCCSDYEDTEE 462
      : : | | :
Db     395 EFHNDNYPDIDD 406

```

RESULT 7

S46584
probable membrane protein YJL094c - yeast (*Saccharomyces cerevisiae*)
N;Alternate names: hypothetical protein J0909
C;Species: *Saccharomyces cerevisiae*
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C;Accession: S46584; S56871; S47057
R;Miosga, T.; Witzel, A.; Zimmermann, F.K.
Yeast 10, 965-973, 1994
A;Title: Sequence and function analysis of a 9.46 kb fragment of *Saccharomyces cerevisiae* chromosome X.
A;Reference number: S46584; MUID:95076716; PMID:7985424
A;Accession: S46584
A;Molecule type: DNA
A;Residues: 1-873
A;Cross-references: UNIPROT:P40309; UNIPARC:UPI000013B5C4; EMBL:X77087; NID:g521093; PIDN:CAA54359.1; PID:g
A;Note: the authors translated the codon TCC for residue 645 as Trp
R;Miosga, T.; Schaaff-Gerstenschlaeger, I.; Baur, A.; Boles, E.; Chalwatzis, N.; Fournier, C.; Schmitt, S.;
submitted to the Protein Sequence Database, September 1995
A;Reference number: S56855
A;Accession: S56871
A;Molecule type: DNA
A;Residues: 1-873
A;Cross-references: UNIPARC:UPI000013B5C4; EMBL:Z49369; NID:g1008267; PID:g1008268; MIPS:YJL094c
C;Genetics:
A;Gene: SGD:KHA1
A;Cross-references: SGD:S0003630
A;Map position: 10L
C;Keywords: transmembrane protein

Query Match 2.2%; Score 110.5; DB 2; Length 873;
Best Local Similarity 21.4%; Pred. No. 0.86;
Matches 69; Conservative 56; Mismatches 106; Indels 91; Gaps 16;

```

Qy      503 VIVVMVAVVVMCLVSIILYRAI--MAIVVSRSGNTLLA-----W-----ASRIASL 547
      | : | : | | : : | : : : | : | : | | : : | : : |
Db      157 VFMVFIASISISVTAFFVLRCRILNELRLIKDRAGIVVLAAGIINDIMGWILLALSIISSA 216

Qy      548 TGSVNVLVFILILS----KIYVSLAHVLTRWEMHRT----QTKFEDAFTLKVFIFQFVNF 599
      | | | | : | : : : | | | : | | : : | : : | : :
Db      217 EGSVPNTVYILLITFAWFLIYFFPLKYLRLRWVLRTHELDRSKPSPLATMCILFIMFISA 276

Qy      600 YSS-----PVYIAFFKGRFVGYPGNYHTLFGVRNEEC-----AAGGCLIELAQ- 642
      | : | : | | | | : | : | : | : | : | : | : |
Db      277 YFTDIIGVHPIFGAFIAGLVVPRDDHYVVKLTERMEDIPNIVFIPIYFAVAGLNVDLTLL 336

Qy      643 -----ELLVIMVGKQVINNMQEVLPKLG-WWQKFLRLSKKRKAGASAGASQGP 691
      : | : : : | : | | : : : : : | : |
Db      337 NEGRDWGVVFATIGIAIFTKIIISG---TLTAKLTGLFWRE-----ATAAGV---- 379

Qy      692 WEDDYELVPCEGLFD-EYLEMVLQFGFVT-----IFVAACPLAPLFALLNNWVEIRLDAR 745
      | : | : : | : | : : : | : | : : | : | :
Db      380 -----LMSCKGIVEIVVLTVGLNAGIISRKIFGMFV-----LMALVSTFVTTPLTQL 426

Qy      746 KVFCEY----RRPVAERAQDIG 763
      : | : : | : |
Db      427 VYPDSYRDGVRKSLSTPAEDDG 448

```

RESULT 8 T00487

probable potassium transport protein F19I3.29 - *Arabidopsis thaliana*
C;Species: *Arabidopsis thaliana* (mouse-ear cress)
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004
C;Accession: T00487; B84764
R;Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul, S.; Mason, T.M.;
submitted to the EMBL Data Library, April 1998
A;Description: *Arabidopsis thaliana* chromosome II BAC F19I3 genomic sequence.
A;Reference number: Z14160
A;Accession: T00487
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-792
A;Cross-references: UNIPROT:O64769; UNIPARC:UPI00000485F4; EMBL:AC004238; NID:g3033373; PID:g3033401
A;Experimental source: cultivar Columbia
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; Mason, T.M.; Bowman
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.

A;Reference number: A84420; MUIID:20083487; PMID:10617197
A;Accession: B84764
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-792
A;Cross-references: UNIPARC:UPI00000485F4; GB:AE002093; NID:g3033401; PIDN:AAC12845.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g35060; F19I3.29
A;Map position: 2
A;Introns: 50/3; 126/1; 208/1; 225/1; 312/1; 368/1; 627/2
C;Superfamily: barley probable potassium transport protein HAK1

```

RESULT 9
A45573
genome polyprotein - hepatitis C virus (strain JT)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructural protein NS3); m
C;Species: hepatitis C virus
C;Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 09-Jul-2004
C;Accession: A45573
R;Tanaka, T.; Kato, N.; Nakagawa, M.; Ootsuyama, Y.; Cho, M.J.; Nakazawa, T.; Hijikata, M.; Ishimura, Y.; S
Virus Res. 23, 39-53, 1992
A;Title: Molecular cloning of hepatitis C virus genome from a single Japanese carrier: sequence variation w
A;Reference number: A45573; MUID:92295714; PMID:1318627
A;Accession: A45573
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-3010

```

A;Cross-references: UNIPROT:Q00269; UNIPARC:UPI0000131E29; GB:D11168; GB:D01171; NID:g221612; PIDN:BAA01943
A;Experimental source: HCV-JT
A;Note: sequence extracted from NCBI backbone (NCBIN:106206, NCBI:106207)
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: ATP; glycoprotein; hydrolase; nucleotide binding; P-loop; polyprotein; serine proteinase; trans
F;2-115/Product: capsid protein C #status predicted
F;116-191/Product: envelope protein M #status predicted
F;192-389/Product: major envelope protein E #status predicted
F;390-729/Product: nonstructural protein NS1 #status predicted
F;730-1006/Product: nonstructural protein NS2 #status predicted
F;1007-1615/Product: hepatitis C virus polyprotein #status predicted
F;1230-1237/Region: nucleotide-binding motif A (P-loop)
F;1312-1317/Region: nucleotide-binding motif B
F;1316-1319/Region: DEXH motif
F;1616-1862/Product: nonstructural protein NS4a #status predicted
F;1863-2013/Product: nonstructural protein NS4b #status predicted
F;2014-3010/Product: nonstructural protein NS5 #status predicted

Query Match 2.2%; Score 108; DB 1; Length 3010;
Best Local Similarity 20.4%; Pred. No. 7.4;
Matches 160; Conservative 86; Mismatches 276; Indels 262; Gaps 41;

```

Qy      128 SAARDRTDMHRTWRETFLDNLRAGLCVDQ-----QDVQDGNNTTVHYALLSASWAVLCYY 182
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      314 SGHRMAWDMMMNWSPT-----TALVVSQLLRIPQAVVDMVAGAHWGVL----AGLAYY 362

Qy      183 AEDLRKLKPLQELPNQASNWSAGLLAWL---GIPNVLLEVVPDPPEYYSCRFRVKNLPR 239
      : | | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      363 -----SMVGNWAKVLIVMLLFAVGVDGVT-----YTTG----- 389

Qy      240 FLGSDNQDT-----FFTSTKRHQILFEILAKTPYGHEKKNLLGIHQLLAEGVLSAAFLPH 294
      | | : | | | | | : | : : | : | : | : | : | : | : | : | : |
Db      390 --GSQARHTQSVTSFFTQGPQRI--QLINTNGSWHINRTALNCNESLNTGFFAALFYAH 445

Qy      295 DGPFKTPPE-----GP---QAPR-LNQRQVLFQHWA----- 321
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      446 KFNSSGCPERMASCSSIDKFAQGWGPITYTEPRDLQRPYCW-HYAPRCQGIVPASQVCG 504

Qy      322 -----RWGK-----WNKYQPLDHVRRYFGEKVLYFAWL- 350
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      505 PVYCFTSPVVVGTTDRSGAPTYNWGANETDVLLNNTTRPPQ--GNWFG-----CTWMN 556

Qy      351 --GFYTGWLLPAAVVG----TLVFLVGCFLVFSDIPTQELCGSKDSFEMCPLCL-DCPF 402
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      557 STGFTKTCGGPPCNIGGVGNLTLCPTDCFRKHPE-ATYTKCGSGP--WLTPRCIVDYPY 613

Qy      403 WLLSSACALAQAGRLFDHGGTVFFSLFMALWAVLLLEYWKRKSATLAYRWDCSDYEDTEE 462
      | | : | | | : | | | : | : | | | | | | | | | | | | | | | |
Db      614 -----RLWHYPCTVNFTIFKVRMYVGVEH--RLSA--ACNWTGRGERCDLED 656

Qy      463 RPRPQ-----FAASAPMTAPNPITGEDEPYFPERSRARRMLAGSVVIVVMVAVVMCLVS 517
      | | : | : | | | | | | | | | | | | | | | | | | | | | | | |
Db      657 RDRSELSPLLLSTTEWQTLPCSFT-----TLPALSTGLIHLHQNIQVQYLYGIGSAVVS 711

Qy      518 IILYRAIMAIIVSRSGNTLLAAWASRIASLTGSSVNLVFIILSKIYVSLAHVLTRWEMH 577
      : : : : : | | | | | | | | : : : : : | : : : : : | : : : :
Db      712 FVIKWEYIVLLF-----LLLADARVCACLW-----MMLLIAQAEAALENLVV---LN 755

Qy      578 RTQTKFEDAFTLKVFIFQFVNFYSSPVYIAFFKGRFVGYPGNYHTLFGV----- 626
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      756 AASLAGADG-----ILSFLVFFCAAWYI---KGRLV--PGAAYALYGVWPLLLLLLALP 804

Qy      627 -----RNEECAAGGCLIELAQELLVIMVG--KQVINNMQEVLI PKLKGWQKFLRSK 677
      | : | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      805 PRAYAMDREMAASCGG-----VVFVGLIILLTSPHYKVFLARLIWWLQYFITRAE 854

Qy      678 KRKAGASAGASQGPWEDDYELVPC---EGLFDEYLEMVLQFGFVTIFVAACPLAPLFAL 733
      : | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      855 AHLVCVWPPLNVRGGRDAIILLTCAAHPILFIDITKLLAILGLMVLQAAITAMPYFVR 914

Qy      734 LNNWVEIRLDARK-----FVCEYRRPVAERAQDIGIWFHILAGLT 773
      : : | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      915 AQGLIRACMLVRKVAGGHYVQMAFMKLAALTGTIVYDHLTPL---QD---WAH---AGLR 965

Qy      774 HLAV 777
      | | |

```

T11129

Query Match 2.2%; Score 106.5; DB 2; Length 519;
Best Local Similarity 17.8%; Pred. No. 0.93;
Matches 94; Conservative 51; Mismatches 167; Indels 215; Gaps 18;

RESULT 11

http://es/ScoreAccessWeb/GetItem.action?AppId=10552515&seqId=775629&ItemName... 11/17/2006

Qy 499 AGSVVIVVMVAVVVMC----LVSIILYRAIMAVVSRSGNTLLAAWASRIASLTGSVVNI 554
 || : | : | : || : || : | : | : | : | : | : | : | :
 Db 294 AGVITFVGALLVGLCDVGRKKLILISRMVLICSWPSFWLLVNYPS-----PGMLLTV 348
 Qy 555 VFILI 559
 ||:::
 Db 349 VFVMV 353

RESULT 13

JC1346

dopamine beta-monooxygenase (EC 1.14.17.1) precursor - mouse

C;Species: Mus musculus (house mouse)

C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004

C;Accession: JC1346

R;Nakano, T.; Kobayashi, K.; Saito, S.; Fujita, K.; Nagatsu, T.

Biochem. Biophys. Res. Commun. 189, 590-599, 1992

A;Title: Mouse dopamine beta-hydroxylase: primary structure deduced from the cDNA sequence and exon/intron

A;Reference number: JC1346; MUID:93080618; PMID:1280432

A;Accession: JC1346

A;Molecule type: mRNA

A;Residues: 1-621

A;Cross-references: UNIPROT:Q64237; UNIPARC:UPI0000029950; GB:S50200; NID:g260872; PIDN:AAB24330.1; PID:g26

C;Comment: This enzyme catalyzes the hydroxylation of dopamine to norepinephrine.

C;Genetics:

A;Introns: 117/3; 166/3; 252/3; 346/1; 401/3; 449/3; 462/3; 482/3; 525/2; 578/3

C;Keywords: catecholamine biosynthesis; copper; glycoprotein; monooxygenase; oxidoreductase; phosphoprotein

F;1-43/Domain: (or 1-46) signal sequence #status predicted

F;44-621/Product: (or 47-621) dopamine beta-monooxygenase #status predicted

F;300-523/Domain: peptidylglycine monooxygenase I homology

F;68,188,476,570/Binding site: carbohydrate (Asn) (covalent) #status predicted

F;350,528/Binding site: phosphate (Ser) (covalent) (by calmodulin-dependent kinase II) #status predicted

Query Match 2.1%; Score 105; DB 2; Length 621;
 Best Local Similarity 20.2%; Pred. No. 1.6;
 Matches 113; Conservative 61; Mismatches 187; Indels 198; Gaps 29;

Qy 110 ADFVLVWEE-----DLKLDQQD---SAARDRTDMHRTWRETFLDNLRA 150
 || : | : | : || : || | | : | : | : | : | :
 Db 102 ADLIMLWSDGDRAYFADAWSDRKGQIHLDSQQDYQLLQAQRTDGLSLLFKRPF----- 155
 Qy 151 AGLCVDQQDVQDGNNTTVH--YALLSASWAVLCYYAEDLRKLPLQELPNQASNWSAGLLA 208
 : | : | : || | : | : | | | : | | | : | | | :
 Db 156 --VTCDPKDYVIEDDTVHLVYGILEE-----PFQSL--EAINTS----- 190
 Qy 209 WLGIPIVNVLEV-----VPDVPPEYYSCRFRVKNLPRFLGSDNQDTFF-----TS 252
 | : | : | : | : | : | : | : | : | : | : | : | :
 Db 191 --GLHTGLLRVQLLKSEVPTPSMPEVDQTMDIRA---PDILIPDNEQTYWCYITELPPRF 245
 Qy 253 TKRHQILFEILAKTPYGHEKKNLLGIHQLLAEGVLSAAFPLHDGP--FKTPPEGPQAPRL 310
 : | : | : | : | : | : | : | : | : | : | : | : | :
 Db 246 PRHHIIMYEAIV-TEGNEALVHMEVFQCAAE---SEDFPQFNGPCDSKMKPD-----RL 296
 Qy 311 NQRQVLFOHWARWGKWKYQPLD-----HVVRYFGEK---VAL 345
 | : : | : | : | : | : | : | : | : | : | : | :
 Db 297 NYCRHVLAAWALGAK-AFYYPKEAGVPFGGPGSSPFLRLEVHYHNPRIQGRQDSSGIRL 355
 Qy 346 -YFAWLGFYTGWLLPAAVVGTLVFLVGCFLVFSDIPTQELCGSKDSFEMCPLCLD-CPFW 403
 | | | | : | : | : | : | : | : | : | : | : | :
 Db 356 PYTATLRRYDAGIMELGLVYTPLMA-----IPPQE-----TAFVLTGYCTDKCTQM 401
 Qy 404 LLSSACALAQAGRLFDHGGTVFFSLFMALWAVLLLEYWKRKSATLAYRWDCSDYEDTEER 463
 | : | : | : | : | : | : | : | : | : | : | : | :
 Db 402 ALQDSGIHIFASQLHTH-----LTGRKVTVLAR-----DGQER 435
 Qy 464 PRPQFAASAPMTAPNPITGEDEPYFPERSRARRMLAGSVVIVVMVAVVVMCLVSIILYRA 523
 | | | | | | | : | : | : | : | : | : | :
 Db 436 KE-----VNRDNHYSF-HFREIRMLKKVTVYPGDVLITSC----- 470
 Qy 524 IMAIVSRSGNTLLAAWASRIASLTGSVVNLVFILILSKIYVSLAHVLRWEMHRTQTKF 583
 : || : | : | : | : | : | : | : | : | :
 Db 471 ----TYNTENKTL-----ATVGG-----FGILEEMCVNYVHYYPQTELELCKSAV 511
 Qy 584 EDAFTLKVFIFQFVNFYSS 602
 : | | | | | : | : | : | : | : | : | :
 Db 512 DDGFLQK--YFHMVNRFS 528

RESULT 14

H82555

c-type cytochrome biogenesis membrane protein XF2460 [imported] - Xylella fastidiosa (strain 9a5c)

C;Species: Xylella fastidiosa

C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004

C;Accession: H82555

R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing and Analysis, Nature 406, 151-157, 2000

A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.

A;Reference number: A82515; MUID:20365717; PMID:10910347

A;Note: for a complete list of authors see reference number A59328 below

A;Accession: H82555

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-646

A;Cross-references: UNIPROT:Q9PAN5; UNIPARC:UPI00000C2A60; GB:AE004054; GB:AE003849; NID:g9107645; PIDN:AAF

A;Experimental source: strain 9a5c

R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Alves, L.M.C.; Araya submitted to GenBank, June 2000

A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohme, M.; Furlan, L.R.

A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.; Monteiro-Vitorello

A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveira, J.F.; Silvestri,

A;Reference number: A59328

A;Contents: annotation

C;Genetics:

A;Gene: XF2460

C;Superfamily: nrfE protein

Query Match 2.1%; Score 104; DB 2; Length 646;
Best Local Similarity 18.9%; Pred. No. 2;
Matches 95; Conservative 46; Mismatches 152; Indels 210; Gaps 19;

```

Qy      167 VHYALLSASWAVLC-----YYAEDLRLKLPLQELPNQASNWSA---GLLAWLGI 212
      | :||: ::||| | ||: || | :|| | |||
Db      45 VQLSLLAGAFALTYAFLGNDFSQYVAENSHSLP--TLYRSTAVWGAHEGSLLW--- 99

Qy      213 PNVLLEVPDPVPEYYSCRFRVKNLPRFLGSDNQDTFFTSTKRHQILFEILAKTPYGHEK 272
      ||| : | | : | :
Db      100 --VLL-----LAGWTASVALRSHTLPATLSA----- 123

Qy      273 KNLLGIHQLLAEGVLSAAFPLHDGPF---KTPPEGQAPRLNQRQVLFQH----- 319
      :||: ||| | | | | | | | | | | | | | |
Db      124 -RILGVLGLIALGFL-ALILFTSNPFARLLPAVPEGNDLNPLLQDPGMIVHPPLLYAGYI 181

Qy      320 -----WARW-----GKWNKYQPLDHVRRYFG 340
      | | | | | | | | | | | | | |
Db      182 GFAVPFAFAVAVLLEGRIDPTWLRWSRPWHTAWALLTLGIALGSWWAYYELGWGGWWFW 241

Qy      341 EKV--ALYFAWL-----GFYTGWLLPAAVVGTLLVFLVGCFLVFSDIPTQE 383
      : | | : || | : | | | : | : | | | : |
Db      242 DPVENASFMPWLIGVALIHSQAITDKRGSFTHWILLAITAFALALLGTFLVRSSVLT-- 299

Qy      384 LCGSKDSFEMCPLCLDCPFWLLSSACALAQARLFDHGGTVFFSLFMALWAVLLLEYWKR 443
      | :| | : | : | : | |
Db      300 ---SVHFAADPV-----RGAFILLIIFTLIGGALLL----- 328

Qy      444 KSATLAYRWDCSDYEDTEERPRPQFAASAPMTAPNPITGEDEPYFPERSRARMLAGSVV 503
      :| || | : | | | :| :| :|
Db      329 -----YARRAPQL--TPVTINMQQRFTPVSRETILLNLL 362

Qy      504 IVVMVAVVMCLVSIILYRAIMAVVSRSGNTLLAAWASRIA----SLTGSVNVLVFILI 559
      : | :| : | | | : | | : | :|
Db      363 LTCACAMVLL-----GTLYPLLADALALGQLSVGPPYFGPLFTLL 402

Qy      560 LSKIYVSL-AHVLTRWEMHRTQT 581
      : : | | ||| : |
Db      403 MTPLIVLLPLGPFTRWQREHPST 425

```

RESULT 15

JQ2034

RNA-directed RNA polymerase (EC 2.7.7.48) - beet cryptic virus 3

C;Species: beet cryptic virus 3

C;Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 09-Jul-2004

C;Accession: JQ2034
R;Xie, W.S.; Antoniow, J.F.; White, R.F.
J. Gen. Virol. 74, 1467-1470, 1993
A;Title: Nucleotide sequence of beet cryptic virus 3 dsRNA2 which encodes a putative RNA-dependent RNA poly
A;Reference number: JQ2034; MUID:93329401; PMID:8336129
A;Accession: JQ2034
A;Molecule type: mRNA
A;Residues: 1-478
A;Cross-references: UNIPROT:Q86632; UNIPARC:UPI00000EE5F9; GB:S63913; NID:g407557; PIDN:AAB27624.1; PID:g40
C;Genetics:
A;Map position: segment RNA2
C;Keywords: nucleotidyltransferase; reverse transcriptase

Query Match 2.1%; Score 103.5; DB 2; Length 478;
Best Local Similarity 20.7%; Pred. No. 1.5;
Matches 53; Conservative 34; Mismatches 96; Indels 73; Gaps 11;

```

Qy      59 RAQEEDSTVLIDVSPPEAEKRGSYGSTAHASEPGGQAAACRAGSPAKPRIADFVLVWEE 118
      :|: | :| | | :| | :| | :| | :| | :| |
Db      109 KARAFDVNTELDKVPYEQSSSAGYGYRSHKGPPEGGE--THMRAISRVKPTLMTAIRPDEE 166

Qy      119 -----DLKLD RQQD SAARDRTDMHRTWRETFLDNLRAAGLCVDQDQVQDGNNTTV 167
      | : | | : | :| : | | |
Db      167 GPEYTILESVPDIGYTTRTQLADLREKTKVRGVWGRAF----- 203

Qy      168 HYALLSASWA-----VLCYYAEDLRLKLP--LQELPNQASNWSAGLLAWLGIPN 214
      || | : : | : | :| :| | :| :| | :| | :
Db      204 HYILIEGTAARPLENFM LGTTFMHIGSDPQLSVPRILHQMKGREGSKWLYA-LDWSSFDS 262

Qy      215 VLEVPDPVPPEYYSCFRVKNLPRFLGSDNQDTFFTSTKRHQILFEILAKTPYGHEK-- 272
      | :| | | | : | :| :| | | :| :| |
Db      263 -----SVTRFEINCAF--NLLKERIEFPNEET-----ELAFE-LSRILFKHKKLA 304

Qy      273 ---KNLLGIHQLLAEG 285
      | : || : : |
Db      305 APDGNIYMIHKGIPSG 320

```

Search completed: October 27, 2006, 20:29:40
Job time : 56 secs